DYNAMICS OF INTACT PROVIRUSES IN PERINATAL HIV-1 INFECTIONS

by

Priya Khetan

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science.

Baltimore, Maryland
May 2021

Abstract

Human Immunodeficiency Virus (HIV) has impacted the lives of nearly 76 million and about a million people become infected each year even with the use of current preventative strategies and antiretroviral therapy (ART). An ART-free remission/cure strategy is needed.

The HIV-1 reservoir poses a major hurdle to developing an HIV cure as it fosters HIV-1 persistence through viral quiescence. Measuring the true size of the reservoir is complex owing to dominance of defective HIV-1 proviruses, making it harder to specifically target therapies towards the intact proviral reservoir. The assays that are currently used under-or over-estimate the size of the reservoir. A recently established PCR assay, Intact proviral DNA assay (IPDA) was found to differentiate between intact and defective HIV-1 proviruses. It has helped in studying the reservoir in adults, but no extensive study using the IPDA have been performed in children.

Objective: To apply the IPDA to study the dynamics of the HIV-1 proviral reservoir in children and inform strategies for ART-free remission.

Methods: IPDA was optimized for use on the limited number of cells that are generally available from children, by determining the best type of control sample, the optimal DNA isolation technique and genomic DNA input amount to analyze each sample. The IPDA was applied to study the composition and the dynamics of the different DNA species in the HIV-1 reservoir in participants from a previously well characterized cohort (PHACS-AMP) in whom a

conventional total HIV-1 DNA assay had been used to quantify HIV-1 reservoirs.

Results: 25 PHACS-AMP participants were eligible for this study. The distribution of intact HIV-1 proviruses/million PBMCs was lower in those that achieved virologic suppression at <1 year of age compared to those that achieved virologic suppression (VS) between 1-5 years of age. In contrast the defective HIV- 1 proviruses demonstrated an expansion in size regardless of age at VS.

Conclusion: Early effective ART decreases the size of the reservoir in children with perinatal HIV-1 infection which is mainly due to the decline in the intact HIV-1 proviral reservoir.

Keywords: perinatal HIV-1 infection, intact proviral reservoir, early ART, early virologic suppression

Primary reader and advisor: Deborah Persaud

Secondary reader: Joseph B. Margolick

Acknowledgments

First and foremost, I would like to thank my PI and mentor Dr. Deborah Persaud for giving me the wonderful opportunity to work in her laboratory and on this great project. Debbie, you have been a phenomenal mentor, always encouraging me to improve and always looking out for my success. I could not have asked for a better mentor and supporter of my career goals. Thank you so much for all the support, constructive criticism, advice and mentoring over the past 1.5 years. I am grateful to have started my research career in your laboratory.

Next, I would like to thank Dr. Adit Dhummakupt (Tang) for his guidance and training throughout my project, and for asking me the difficult questions which made me think through the work. Thank you for your help with the Gentra Puregene isolation, for providing the primer-probe images for my research and helping me navigate the world of writing of a thesis. Thank you, Joseph Szewczyk (Joe), for all your help with the PHACS samples and the motivation to get through my tough days. Thank you, Laura, for teaching me the ropes of DNA isolation and ddPCR, and for being the ray of sunshine in the dark. Thank you, Ya hui, for all your help with the WGA and the near full-length PCRs. Laura, Joe, Ya hui and Tang, you have all taught me different techniques and skills in my time at the lab and I just want to thank all of you from the bottom of my heart for that.

I would like to thank Dr. Joseph Margolick, for acting as my secondary advisor and reader. Thank you for pushing me to think harder and get down to the basics so that I could understand the topic in more detail.

I would like to express my gratitude for Dr. Kunjal Patel and Brad Karalius and thank them for all their help with analyzing the data and helping me push my thesis forward. A massive thank you to Dr. R.B. Jones and Dr. Francesco Simonetti for providing primers and probes for the IPDA and sequencing, respectively.

I am grateful for the support and infrastructure the Department of Molecular Microbiology and Immunology has given me to complete my ScM degree.

A major thanks to my friends, Kelsey, Whitney, Jade, Mihra, Harshini, Malavika, Ketki, Zak, Stephanie and Shivani for being there for me over the past two years, and for pushing me to complete my goals both personal and professional. Thank you to Last Chance Animal Rescue for giving me, my dog, my lifeline, Miss Spider. Miss Spider, without you I would not have been able to reach this point in my degree.

Many thanks to all my mentors from my undergraduate institution for supporting my ideas, pushing me forward and having faith in me.

Lastly, I want to thank my parents and my sister for supporting me through all my endeavors.

Dedication

This thesis is dedicated to my parents, Sonal and Rajesh Khetan for their support; my sister, Vidhi Khetan, for introducing me to viruses and supporting me throughout my career; to my partner in everything, Luke J Cox, thank you for your love, endless support, and words of encouragement. Lastly, it is dedicated to all the children at Ashray who inspired me to learn more about combatting HIV-1.

Table of Contents

Abstract……ii

Acknowledgments……..iv

Dedication……vi

List of Tables……ix

List of Figures……...x

Chapter I: Introduction……1

i. History of HIV-1……1
ii. HIV-1 Biology & Pediatric HIV-1 Infection……2
iii. Antiretroviral therapy……...10
iv. Definitions……..11
v. Cure Strategies/Long-term ART-free remission……13
vi. Measuring the reservoir ……...18
vii. Reservoir studies in adults using the IPDA……29
viii. Reservoir studies in pediatrics……38
ix. Study aims……..42

Chapter II: Material and Methods……...43

i. Establishing an Internal Standard for the IPDA……43
ii. Preparation of JLAT Standards……45
iii. Isolation of Genomic DNA from JLAT/PBMC standards……48
iv. Intact Proviral DNA Assay……51
v. Whole Genome Amplification and Primer Sequencing……60
vi. HIV-1 Parameters and Characteristics for the PHACS-AMP Cohort…68

vii. Statistical Analysis……...72

Chapter III: Results……75

i. Optimizing the Intact Proviral DNA Assay……75

ii. Applying the IPDA to PHACS samples……..80

iii. IPDA vs Pol-LTR on PHACS samples……102

iv. Determination of primer-probe mismatch……103

Chapter IV: Discussion……...112

i. Optimization of the IPDA……112

ii. Applying the IPDA to children living with perinatal HIV-1 in the PHACS-AMP cohort……118

iii. Conclusions and future directions……129

References……..132

Appendix……..145

CV Priya Khetan……..150

List of Tables

Table 1: JLAT Standard Preparation……47

Table 2: Primer Probe Sequences for IPDA……53

Table 3 (A): Buffer preparation for Whole Genome Amplification……61

Table 3 (B): Preparation of Master Mix for Whole Genome Amplification……61

Table 4: Primers for Psi and Env PCR……62

Table 5 (A): Preparation of Master Mix for Psi-Env Outer PCR……..63

Table 5 (B): Preparation of Master Mix for Psi-Env Nested PCR……64

Table 6: Primers for the near-full length Four fragment PCR……..65

Table 7: Preparation of Master Mix for Four fragment outer PCR……..65

Table 8: Preparation of Master Mix for Four fragment inner PCR……...67

Table 9: HIV-1 parameters and characteristics, by age at virologic suppression……70

Table 10: Medians of HIV-1 DNA species segregated based on duration of virologic suppression……95

List of Figures

Figure 1. HIV-1 & HIV-1 Lifecycle……8

Figure 2. Formation and composition of the HIV-1 reservoir……19

Figure 3. IPDA primer binding sites……..29

Figure 4. 2D amplitude in Bio-Rad Quantasoft showing droplets……57

Figure 5. Intact HIV-1 DNA copies/million with different DNA input amounts……..77

Figure 6. Range of % Unsheared observed over a series of runs……79

Figure 7. % Unsheared observed from DNA isolated using two different methods and kits……..81

Figure 8. % Unsheared for the cohort……83

Figure 9. Total cells analyzed and cells analyzed/well for the cohort…….85

Figure 10. Distribution of PBMC-associated intact HIV-1 DNA……86

Figure 11. Distribution of PBMC-associated 5'-deleted (defective) HIV-1 DNA……88

Figure 12. Distribution of PBMC-associated 3'-deleted/hypermutated (defective) HIV-1 DNA……89

Figure 13. Distribution of PBMC-associated total (intact and defective) HIV-1 DNA……90

Figure 14. Distribution of PBMC-associated intact HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant)……..92

Figure 15. Distribution of PBMC-associated 5'-deleted (defective) HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant)……..93

Figure 16. Distribution of PBMC-associated 3'-deleted/hypermutated (defective) HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant)……..……97

Figure 17. Distribution of PBMC-associated total (intact + defective) HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant)……98

Figure 18. Distribution of various HIV-1 DNA species as a function of duration of Virologic Suppression……..99

Figure 19. Proportion of total HIV-1 DNA that was intact……100

Figure 20. Correlations between Total HIV-1 DNA Copies/Million PBMCs detected by POL-LTR assay vs HIV-1 DNA Copies/Million PBMCs of all species by IPDA……104

Figure 21. Primer-Probe mismatch in participant samples……..106

Figure 22. Gel Electrophoresis of Nested PCR for u5gag and env regions……....107

Figure 23. Gel Electrophoresis of Repeated Nested PCR for u5gag and env regions……108

Figure 24. Near full-length four fragment PCR Gel Electrophoresis……109

Figure 25. Sanger Sequencing and alignment results for env region mismatch……110

Figure 26. Sanger Sequencing and alignment results for u5gag region mismatch……..111

Supplementary Figure 1: Individual plots for participants whose total HIV-1 DNA copies/million PBMCs were in flux over time on ART……145

Supplementary Figure 2: Individual plots for participants whose total HIV-1 DNA declined over time on ART……146

Supplementary Figure 3: Individual plots for participants showing stable total HIV-1 DNA copies/million PBMCs over time on ART……147

Supplementary Figure 4: Individual plots for participants with only one time point……148

Supplementary Figure 5: Individual plots of participants that need further testing……149

I. Introduction

History of HIV-1

The human immunodeficiency virus type 1 (HIV-1) is a retrovirus belonging to the family of lentiviruses that has a diploid single stranded RNA genome. In 1981, Acquired Immunodeficiency Syndrome (AIDS), a new disease prevalent amongst men who have sex with men (MSM), blood transfusion recipients and injection drug users was discovered[1,2]. AIDS was detected in children around 1982[3]. With AIDS, the immune system of the patient is compromised to the point that opportunistic infections that are not usually harmful, become deadly. After much research and collaboration, HIV-1 (then called HTLV-III) was discovered to be the causative agent of AIDS in 1983[1,4,5]. Later, in 1986, a similar virus causing immunological symptoms in people in West Africa was discovered and termed HIV-2. It was found that without treatment, HIV-1 has higher mortality rates than HIV-2[6]. The life expectancy, without treatment, for children with HIV-1 was about 2-3 years while adults lived on for a median of about 8-10 years after infection[7,8]. The modes of transmission of HIV-1 in adults are through sexual contact, sharing needles and blood transfusions[1,4,5,9]. In infants, the virus can enter infant cells in-utero, during the delivery, or after birth with breastfeeding[10,11]. Since the discovery of HIV-1, extensive studies have been performed to combat the HIV/AIDS epidemic and prolong lives with antiretroviral treatment (ART)[12-19].

HIV-1 has affected the lives of nearly 76 million people since the start of the epidemic.

Even with the current preventative strategies and antiretroviral therapies, (ART), nearly 1.7 million people were newly infected in 2019 of which about 150,000 were children[20]. HIV-1 continues to pose a public health problem especially in the pediatric population where measures have been taken to diminish mother to child transmission[3,11,21]. ART is a lifelong regimen which needs to be strictly adhered to, for it to be effective. However, it is not easy to adhere to it strictly especially for the pediatric population. Therefore, the field has moved towards finding a cure or ART-free remission for HIV-1. Clearly, ART is not enough to provide a cure therefore extensive studies are ongoing to follow through on this agenda[22-24].

HIV-1 Biology & Pediatric HIV-1 Infection

Biology: HIV is a lentivirus belonging to the family *Retroviridae*. It has a diploid positive strand RNA genome which is 9 kilobases in length. HIV is of two types: HIV-1 and HIV-2. It is a zoonotic disease that originated from cross-species transmission from chimpanzees to humans. HIV-1 is found all across the world and has several clades which are distributed amongst 4 groups: M (main), O, N (non-M and Non-O) and P. HIV-1 group M is the most common and has nine subtypes: A, B, C, D, F, G, H, J & K. It also has 49 circulating recombinant forms (CRFs). Of the nine subtypes, subtype B viruses are most commonly found in Europe, Australia, and North America. All the subtypes of group M can be found in Africa. However, subtype C viruses are common worldwide while clade A viruses circulate mainly in Eastern and Central African countries.

As for the other groups of HIV-1, group O has been isolated from patients in Cameroon, Equatorial Guinea, and Gabon, while group N and P have been isolated from patients in Cameroon as well.

These different subtypes/clades of HIV-1 could have arisen as a result of error-prone RNA polymerase activity or genetic recombination as part of reverse transcription of the HIV-1 genome[6,25].

HIV-1 has an affinity for infecting cells that express CD4. T cells and some monocyte-derived macrophages (MDM) express CD4 as a cell surface receptor. Therefore, HIV is characterized by its tropism for T cells and MDMs as T-Cell tropic and M-tropic respectively. Viruses that infect both T cells and MDMs have also been observed and are termed as dual tropic. For successful entry, HIV usually uses a co-receptor, CCR5 (R5 virus) or CXCR4 (X4 virus), adding another level of tropism for the virus[6,26].

The core of HIV-1, i.e., nucleocapsid, genome and capsid, is cone shaped and cylindrical (Figure 1 A). The 9 kilobase genome encodes several genes and has 10 open reading frames (ORFs). Of the 10 open reading frames, only 3 genes are common across all retroviruses: gag, env and pol. The gag and env genes encode structural proteins while the pol gene encodes enzymes that are important for viral replication. The remaining genes Vif, Vpr, Tat, Rev, Vpu, and Nef are accessory and have several essential functions in the viral lifecycle[6,25].

The gag and pol genes are translated as a polyprotein and are separated by a single frameshift giving rise to 2 ORFs.

The env gene encodes a protein product, which when cleaved using cellular proteases forms two components: gp120 which forms the main spike protein of HIV-1 and the transmembrane region gp41 which acts as the fusion peptide for membrane fusion during entry. The gag protein consists of the matrix, capsid, late domain (p6, p1, p2) and nucleocapsid proteins.

Within the matrix (MA) protein domain the plasma membrane targeting, membrane binding regions and the nuclear localization signal that allows the viral genome to enter the nucleus can be found. The capsid (C) domain has gag-binding regions which allow gag to multimerize, and the nucleocapsid (NC) domain has zinc finger regions that allow RNA binding to take place through interaction of the four RNA stem loops that make up the core packaging signal, psi (ψ). The late domain has proteins like p6 which are required for budding. The Pol protein contains the enzymes protease, reverse transcriptase, and integrase which are essential for viral infectivity. The viral protease is important for cleaving gag and gag-pol into their components. Reverse transcriptase transcribes the RNA viral genome to viral cDNA, and integrase inserts this cDNA into the host genome. The 5' untranslated region of the viral RNA genome has several secondary structures that can mediate elongation, splicing of the viral transcripts, reverse transcription, packaging and dimerization of viral genome. There are certain regions in this 5' UTR where the t-RNA primer needed for reverse transcription can bind and be packaged into the virion (mediated by the NC domain).

HIV-1 Lifecycle: The virus enters the human host through several routes including but not limited to sexual contact or blood transfusions. In infants, the virus can enter cells in-utero, during delivery, or after birth with breastfeeding. The virus then binds to the CD4 receptor on the T cells or monocyte-derived macrophage (MDM) using the gp120 of the env protein which leads to a conformational change. This change allows the gp120 to bind to a co-receptor, usually either CCR5 or CXCR4, leading to another conformational change. The second conformational change allows the gp41 to be exposed and for fusion to occur at the plasma membrane.

Until recently (Figure 1B), it was believed that once fusion occurred, the viral core would disassemble soon after and the viral pre-integration complex (viral cDNA, integrase, vpr, gag MA p17) would then enter the nucleus and facilitate integration of the viral DNA[6,25,26]. However, it was recently discovered that the viral core remains intact even after fusion and enters the nuclear envelope followed by traversing the nuclear pore complex into the nucleus[27]. Within the intact core, the virus's reverse transcriptase (RT) binds to the HIV-1 RNA and converts it into cDNA (Figure 1B). It has been proposed that the dNTPs for RT enter the core though pores in the capsid[27,28]. Reverse transcription begins when the t-RNA primer binds to the Ru5 region of the HIV-1 RNA template and starts transcribing the RNA to DNA and then stops. The RT then jumps to the u3R region of the template RNA and begins transcribing it to DNA. During this jump, the RT may latch onto the other copy of RNA if it comes across a break or nick in the template strand or if there is more than one virus infecting the same cell (quite rare)[29].

As the reverse transcription proceeds the viral RNA is degraded by the RNAse activity of RT, except RNA near the u3R region which acts as a primer for positive strand synthesis.

Once reverse transcription is complete, the RU5U3R portion of the cDNA molecule is present in duplicate and forms what is known as the long-terminal repeat (LTR). LTRs are important for the integration of viral cDNA into the host genome. After recognizing certain sequences within the 2LTR region, the enzyme integrase is then able to co-linearly insert the HIV-1 DNA into the host genome. This leads to the formation of the proviral reservoir in cells.

The HIV DNA can get inserted into transcriptionally active or inactive genes, and in the opposite orientation to the human gene.

This integrated DNA can then be transcribed to form spliced or unspliced mRNA. After the virus has generated mRNA (multiply spliced), it transcribes the viral mRNA using the host cell ribosomes into viral proteins using the accessory proteins tat and rev. As the cycle moves towards the later stages, more unspliced (full-length genome) or single spliced mRNAs are generated. The unspliced mRNAs are then packaged to form new virion particles. Packing occurs when sequences in the packaging region of the genome generate a kissing loop that allows the two RNA strands to bind to one another non-covalently at the 5’ ends. Individual gag monomers using the Zn finger associate with the other secondary structures in the packaging region of the viral RNA genome. This allows the specific interaction of gag nucleocapsid (NC) with the viral RNA; gag then multimerizes to form a higher-level structure, which is preferred for packaging. With the help of the late domain, p6, the gag binds

to host ESCRT proteins that are needed for assembly and budding. The gag matrix domain allows it to be targeted to the lipid-rich regions of the plasma membrane where the HIV-1 env proteins are also present. As gag binds to the lipid-rich regions, it also multimerizes, leading to lipid concentration in one area, and through complex mechanisms involving the ESCRT host proteins the virion buds out of the cell. Once the virus has budded, the host cell proteins packaged into the virus cleave the gag MA region to form a stable mature structure[6,25,26].

Pediatric HIV-1 Infection: Nearly 150,000 children were newly infected with HIV in 2019[20]. HIV-1 can be transmitted from a positive mother to her child in three ways: in-utero, intra-partum, or post-partum through breast feeding[8,10]. The neonate is said to be infected in-utero if two independent blood samples taken within 48 hours of birth show positive results for HIV-1 DNA or RNA. A neonate is said to be infected perinatally (intra-partum or post-partum) if samples of blood taken within a week of birth are negative but 2 or more subsequent samples obtained from blood drawn on different visits are positive for HIV-1 DNA.[8,10,11]

If the infection is transmitted in-utero, the infection progresses more rapidly than if it is transmitted at the other two times. The reasons stated for more rapid progression of the infection are the immunology of the neonate and ability to treat sooner with respect to timing of infection especially in cases of peri-partum infections[8,30]

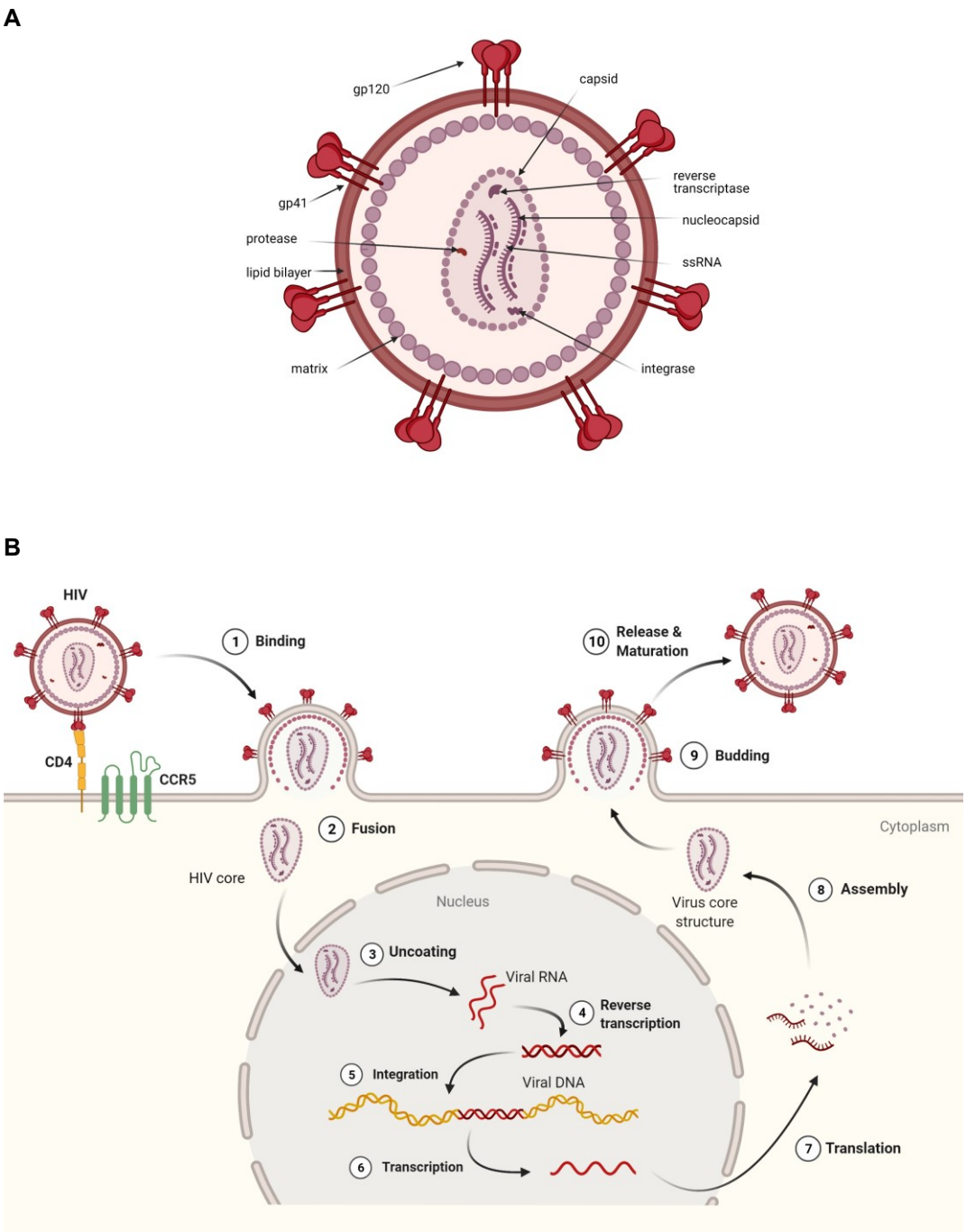


Figure 1. HIV-1 & HIV-1 Lifecycle

A: The HIV-1 virion conical core (NC, RNA and CA) is encased by the matrix followed by a lipid bilayer. The spike protein on the surface gp120 is embedded in the bilayer and is used for attachment to the cell[6, 26]. B. The life cycle of HIV-1 follows the steps of attachment, fusion[27], uncoating of the core in the nucleus vs the original belief of uncoating right after fusion, reverse transcription, integration, viral protein generation, assembly, and release of the virion[6, 25, 26].

Figures created with Bio-render.com.

The immunologic environment of the fetus in-utero is tolerogenic owing to exposure to maternal antigens through the placenta[30]. The inflammatory responses are controlled by high levels of transforming growth factor-beta (TGF-β). TGF-β pushes the naïve CD4 T cells into a CD4 Treg lineage[31]. In addition, innate immune cells respond to TLR stimulants with the release of cytokines that push the differentiation of naïve CD4 T cells in the direction of CD4 Th17 cells and Th2 cells rather than Th1 cells. This gives rise to a predominantly anti-inflammatory environment[30,32-34].

In the case of HIV-1 infection, this tolerogenic environment can interfere with the HIV-1 latency. The immune repertoire circulating in the blood of neonates is dominated by naïve CD4 T cells (low to no expression of CCR5) while the intestinal epithelium consists of a large proportion of CD4 T cells expressing CCR5[35,36]. Some of these intestinal T cells are CD4+ CCR5+ CCR6+ T cells (Th17) and are highly permissive for infection by HIV-1.[37,38] These Th17 cells then decline in number and could leave the patient vulnerable to immune activation as a result of infection by extracellular pathogens or commensals[30]. Immune activation can be caused by response to vaccines, mixed feeding (breast and other forms of feeding the neonate), changes in the microbiota as the neonate grows. Increased immune activation leads to increased HIV replication, paving the way for the formation of a reservoir of HIV-1. Thus, the tolerogenic environment of the neonate can change based on multiple factors paving the way for a more robust infection by HIV-1 [30].

Antiretroviral Therapy

In 1980s, a drug used in cancer treatment, was found to be capable of inhibiting reverse transcriptase, a direct acting dideoxynucleoside reverse transcriptase inhibitor (NRTIs), azidothymidine (AZT) and was used for treating HIV-1 infections[39]. It was after this step that different categories of inhibitors were developed to be combined to bring virus replication under control: non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors, entry and attachment inhibitors and integrase inhibitors. In the late 1990s, combination antiretroviral therapy (now called ART) to potently suppress virus was replication was identified[18,40-47]. Over the past two decades, these new drugs when combined gave rise to highly active antiretroviral therapy (now called ART), improved the outcomes for HIV-1/AIDS in terms of life expectancy and reduced the count of tablets taken per day to 1-3. The transmission of HIV-1 from a HIV positive mother to her child also is reduced with the use of effective ART[3,21].

ART has given people living with HIV and AIDS (PLWHA) benefits such as increased life span and better quality of life, but a patient has to adhere to it strictly for it to be effective in keeping the plasma viral load at undetectable levels and to prevent the selection of drug-resistant strains of HIV-1. Strict adherence to ART is not easy and even with effective ART, the virus cannot be fully eliminated[48-52]. Treatment interruptions give rise to viraemia which is a consequence of the reservoir of HIV-1[12,53]. Even with the field moving towards early effective ART given at the time of diagnosis of HIV-1 infection, the reservoir still persists, albeit smaller compared to when ART is initiated during

chronic infection[54-62]. Hence, there is a need to develop new lines of treatment that can specifically target the reservoir so that people living with HIV/AIDS (PLWHA) can be taken off treatment without rebound of viremia for long periods of time[13].

Definitions

Early effective/suppressive ART: Early effective ART refers to the administration of ART within 3 months of infection, and the achievement of control of viremia in a few weeks post initiation of ART[55,63].

Virologic suppression: Control of viremia such that the plasma viral load (pVL) ≤400 copies/mL for the duration of the study while permitting single measurements of pVL≥400 copies/mL in between measurement of pVL≤400 copies/mL[59,60].

HIV-1 provirus: HIV-1 DNA that is integrated into the genome of an infected cell[6,64].

Viral replication: A process in which virus production leads to infection of new cells (previously not infected), which then produce infectious virus that can infect more non-infected cells[23].

Replication-competent HIV-1 provirus: HIV-1 provirus that is capable of producing and releasing virions that can cause productive infection of cells[23,64-68].

Intact HIV-1 proviral genomes: HIV-1 genomes lacking fatal small and large deletions, insertions, premature stop codons or hypermutations [64,65,69].

Defective HIV-1 proviral genomes: HIV-1 genomes which are not capable of producing infectious viruses. Defective proviruses contain large deletions, spanning one or more regions of the genome, insertions, frameshift mutations, hypermutations mediated by APOBEC3G/F, and mutations that affect viral fitness [69-73].

Latency: A reversible state in which cells with integrated HIV-1 genomes do not produce virions and expression of viral proteins is limited or absent (transcriptionally silent)[23,49,64-68,74].

Cure: There are several definitions of cure in HIV-1 infection.

Sterilizing cure: Eradication of all forms of replication-competent virus[23,65,75].

Functional cure: Sustaining virologic suppression (control of viral replication) without antiretroviral therapy[23,65,75].

ART-free Remission: In the context of perinatal infection, remission is defined as the ability to sustain virologic suppression in the absence of ART, while maintaining normal CD4 levels and immune responses to childhood vaccines[76].

Reservoir: Population of cells or anatomical sites that allow the persistence of replication-competent proviruses for several years even in patients on effective ART. Reservoirs are relatively stable as they are protected from ART or the immune system compared to actively replicating virus. Currently, the reservoir is thought to be found mostly in resting CD4 T cells[77,78].

Resting CD4 T cells can get activated upon presence of stimulus, become effector cells permissive of viral production and gene expression (short-lived) and then revert to the resting phenotype wherein they become long-lived cells again[66]. There is increasing evidence that other cell types can also act as reservoirs[79,80]. For example, macrophages, which are relatively long-lived once infected,[81-83] and other compartments in the body like the CNS,[84,85] can act as reservoirs, however further studies are needed to confirm this with respect to longevity of the cells/tissues [23,66,68].

Latent reservoir : A population of quiescent cells that harbor HIV-1 proviruses that are replication-competent but do not express viral proteins or produce virions while the cells are quiescent. The latent reservoir is not affected by the immune system or ART[23,64-66,68,86].

Active reservoir: Transcriptionally active cells that harbor HIV-1 proviruses capable of producing HIV-1 RNA and proteins, found in persons on effective ART[64,68]

Inducible HIV-1 provirus: Integrated HIV-1 DNA in infected cells which can be stimulated by antigens or activating agents to produce replication-competent virus and express viral proteins. [23,49,64-68]

Cure/Long-term ART-free Remission

After the discovery of HIV as the cause for the fatal disease AIDS, there was an urgent need for therapies[1,4,5]. This need was met with the development of ART beginning with repurposing AZT from cancer treatment to HIV-1 treatment.

Later, combination antiretroviral therapy was developed by using two or more drugs to control the viral lifecycle and suppress viral replication to a point where viremia was below the detection limit of clinical assays[13-16,39,42,43,45-47]. Although effective ART was able to arrest viral replication and control viremia, it was unable to cure HIV-1 infection as evidenced by the rebound of viremia within 2-4 weeks of ART interruption (ATI)[53,87]. The reason for this rebound was found to be replication-competent integrated HIV-1 DNA in quiescent long-lived memory CD4 T cells (latent reservoir). [48,49,52,77,88]. This latent reservoir poses a barrier to cure as it can be reactivated[89]. The latent reservoir is persistent and has limited to no expression of viral proteins while the cells are in a resting state. Owing to this, cells in the latent reservoir cannot be differentiated from non-infected cells and therefore are not subject to immune responses[89]. Studies on people infected with HIV, who were on effective ART for 7 years showed that the latent reservoir decays slowly with a half-life of 44 months which indicates that a person would have to stay on effective ART for at least 73.4 years to fully eradicate one million latently infected cells[50,51,90,91]. Even when ART is initiated early, that is during acute HIV-1 infection, upon cessation of ART viral rebound is observed[53,92-94]. In addition, ART is not easily accessible to everyone, adherence to the regimen can be difficult, and ART has toxicity associated with long-term use[95,96]. In spite of current preventative and treatment strategies, at least 1.7 million new infections of HIV-1 occur worldwide annually [20,97]. The issues with ART have led to the search for a cure for HIV-1 infection beyond using ART[96,98].

Seeding and maintenance of the reservoir

Naïve T cells do not express high levels of cell surface receptors. When the naïve T cell is exposed to an antigen-presenting cell carrying a cognate antigen, cell surface receptors like CD4, CCR5, CD28 etc. are upregulated on the T cell to assist in antigen recognition and elicit an appropriate immune response. When a person is infected with HIV-1, the expression of high levels of cell surface receptors (CD4, CCR5/CXCR4) allows the virus to enter the cell. The activated T cells are also transcriptionally active, permitting viral infection. This favorable environment permits the HIV-1 virus to infect activated T cells i.e., memory T cells that are activated in response to a cognate antigen. At the end of an immune response, the majority of the activated T cell repertoire is cleared but a small fraction transitions into resting memory T cells which are long-lived and allow the virus to persist[99]. However, the average half-life of memory T cells, is approximately 5.5 months, which is shorter than that of the HIV-1 reservoir (44 months)[50,100-102]. Therefore, there must be mechanisms in place that help maintain the reservoir in patients after years on effective ART.

Three such mechanisms have been identified: homeostatic proliferation, antigen driven proliferation, and integration into/in close proximity to genes involved in cell growth. These mechanisms collectively fall under the term clonal expansion[96,98,103]. Homeostatic proliferation is a part of natural T cell homeostasis and occurs when the T cells are exposed to cytokines like interleukin (IL)-7 and IL-15[104-106]. Homeostatic proliferation can lead to increase in cell numbers without expression of the HIV-1 genome, thereby protecting the infected cell from the immune response[107].

Antigen driven proliferation occurs through repeated exposure to cognate antigens[108]. It is possible that this type of proliferation may lead to expression of HIV-1[103].

HIV-1 tends to integrate into active genes[109,110]. Sometimes these integration sites include genes which are associated with cell growth, and can lead to cell proliferation, thereby contributing to the maintenance of the reservoir[111,112].

It was observed that majority of HIV-1 clones were from cells harboring defective proviruses[113]. Recently, it was shown that even cells harboring intact, replication-competent proviruses can expand clonally[114-117].

Latency

There are several mechanisms by which HIV-1 maintains its latency. Studies have found the following mechanisms to play a role in the maintenance of HIV-1 latency[66,67,96,118]:

- Repressive chromatin states: including but not limited to posttranslational modifications(histone + non-histone proteins)[119-122] and DNA methylation[123].
- Preference for sites of integration: HIV-1 prefers to integrate into inter-genic regions of the active genes, where epigenetic silencing may play a role in the maintenance of latency in cells.[96,109,110]
- Low levels of host transcription factors: Low levels of host transcription factors such as NF-kB, NFAT, AP-1 and low levels of viral protein tat can interfere with viral protein production[79,124-127].
- Repressive genetic components: Nucleosomes present in the 5'LTR region of the viral DNA can block viral transcription[128]. In addition, low levels of HIV-

1 tat and another host factor P-TEFb can interfere with the elongation of transcripts leading to formation of truncated proteins[129].

- Orientation of the integrated HIV-1 DNA with respect to the host genome: if the viral genome is integrated in the opposite orientation to that of the host genome, the RNA pol II complexes could collide leading to RNAi, and reduction of RNA transcripts of the host and virus. Alternatively, if the genome is integrated in the same orientation as that of the host, the RNA pol II upstream of the viral genome could disrupt the positioning of the factors required for viral transcription present at the 5'LTR promotor site[130-132].

Strategies for cure/long-term ART-free remission

Several strategies for cure/remission are under study. These include: latency reversal agents which involves the use of mitogens to activate latently infected T cells such that they express HIV-1 and can be eliminated (shock and kill)[133-138]; latency silencing agents (block and lock)[139-141]; the use of broadly neutralizing antibodies to target viral surface proteins on latently infected cells after latency reversal followed by elimination of the infected cell[142,143] and others. The aim of these strategies is to reduce the size of the replication-competent reservoir which poses a barrier to cure[96,98]. To determine if any of these strategies are working especially in the setting of a clinical trial, a simple, efficient, cost-effective, high-throughput assay is required to study the reservoir with respect to intact proviruses. The Intact Proviral DNA assay (IPDA) can study the intact proviruses and determine the decay rate of intact proviruses, thereby proving to be a helpful assay in cure strategies. It is also requires very

few cells and can be used to study several types of tissues and cells[64,65,69,144,145].

Measuring the reservoir

To find a cure, whether functional or sterilizing, it is important to measure the size of the reservoir accurately to observe the effects of the therapies/strategies on it and note whether the strategies are effective in eliminating the reservoir (sterilizing cure) or the replication competent proviruses (functional cure). To achieve this, the method used to measure the reservoir should satisfy a few criteria: The assay/method should ideally measure only the replication-competent proviruses. The assay should be able to work across different subtypes, and mutations should not drastically impact the observed results. It should be sensitive and specific to not deliver false negatives. Ideally, it should not require large specimen volumes or samples[65,68].

- *Assays that measure the concentration of intact or replication-competent proviruses*
  - Culture-based:

Quantitative Viral Outgrowth Assay

The quantitative viral outgrowth assay (QVOA) also known as viral outgrowth assay (VOA), was the first assay used to identify the reservoir in HIV-1 infections[74,77,146]. It is currently considered the gold standard for measuring the reservoir[23]. The principle of QVOA is based on viral gene expression and state of activation of the cell.

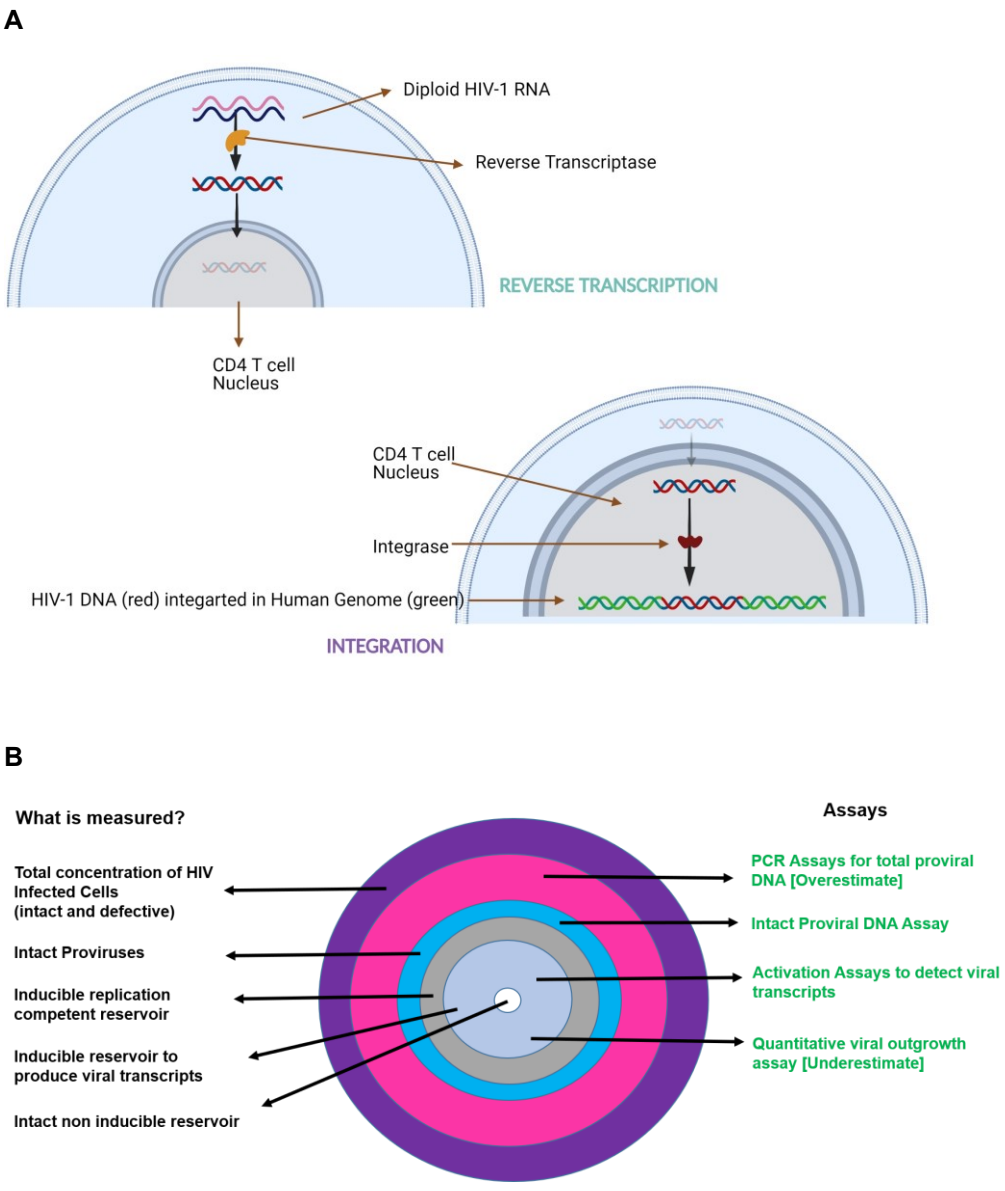


Figure 2. Formation and composition of the HIV-1 reservoir

A: The proviral reservoir is formed by conversion of viral RNA to viral cDNA within the pre-integration complex followed by insertion of the viral cDNA into the host genome and survive in a long-lived quiescent CD4+ T cells

B: The proviral reservoir is complex and consists of different species of HIV-1 DNA represented here as different colored circles which are measured using the different assays summarized below[96,144]. On the left side of the figure, are the different HIV-1 DNA species that constitute the reservoir and on the right are current assays that are used to measure the different DNA species in the reservoir[23, 64, 65, 68, 69, 73, 77, 90, 96, 116, 144, 145, 148, 151, 155, 178].

Figures created with BioRender.com.

In the QVOA, purified CD4 T cells from blood or leukapheresis products of people infected with HIV-1 are mixed with stimulants such as mitogens (phyto-hemagglutinin, histone deacetylase inhibitors, anti-CD3/anti-CD28 antibodies, irradiated PBMCs depleted of CD8 cells[147]) in a limiting dilution to push the integrated provirus out of latency and promote HIV-1 gene expression. To this, activated CD4 T cells from HIV-1-negative donors are added for viral outgrowth. A capsid antigen, p24, which is detected in the supernatant of the culture with ELISA, is used to measure viral outgrowth. Using Poisson's distribution, the infectious units per million (IUPM) CD4 T cells can be calculated (only possible if limiting dilution is used) [23,52,68,74,144,148-151].

The QVOA has been able to detect 0.03 to 3 replication-competent proviruses per million resting CD4 T cells[50,148]. Earlier CD4 lymphoblasts from HIV-1 seronegative individuals were used as target or indicator cells; now, cells lines or PBMCs can be used[152]. The QVOA can detect single infected cells, does not detect defective proviruses and is a reproducible, reliable, minimal estimate of the replication-competent proviral reservoir size. It has been helpful in measuring the stability of the reservoir over time. However, the assay is labor-intensive, expensive, time-consuming, has a long turnaround time and requires a large number of cells to achieve a low limit of detection[64,65,68,74,144,151,153]. It also underestimates the size of the latent reservoir by ~ 25-60-fold as not all replication-competent proviruses are induced in one round of activation[70,73,116]. Modifications have been made to the assay to reduce the duration and improve detection in shorter times such as moving from the ELISA to PCRs for HIV-1 RNA, using the HIV infection permissive cells line MOLT-4 with CCR5 expression and more[152,154,155].

As the proviral reservoir is dominated by defective proviruses, the switch to the use of these sensitive assays is not beneficial because some defective proviruses are capable of producing viral RNAs which will lead to overestimation of the reservoir size [70-73,152,154,156].

These cons make it hard to use the QVOA in large studies such as clinical trials. However, it still remains a prime minimal estimate assay for studying the latent reservoir in HIV-1 infected people [64,65,68,86,144,151].

- PCR- and sequencing- based:

Quantitative PCR

As culture-based assays like QVOA were laborious and complex, newer and simpler approaches to study the latent reservoir were developed, among them was the use of quantitative PCR assays[68,144]. qPCR is an assay that is performed on PBMCs or CD4 T cells to study the HIV-1 reservoir. The primer-probes for this assay usually target conserved regions of genes like POL[153,157]. The results are interpreted by generating a standard curve from plasmid controls and calculating the relative quantities of the HIV-1 DNA copies[144,153,157]. The qPCR assays are sensitive, specific, cost-effective, have short turnaround times, are less complex, need smaller number of cells compared to QVOA, can be performed on different types of tissues/cells (primarily CD4 T cells/PBMCs) and sequencing of the proviruses detected is possible[64,68,144,153,155].

However, qPCR assays also have several disadvantages: Owing to the heterogeneity of the HIV-1 genome, qPCR assays are prone to primer-probe mismatches and it is difficult to identify whether a negative reaction is a true negative or a result of primer-probe mismatches[65,144,153,155].

Upon comparison with QVOA, it was found that the frequency of infected cells detected by qPCR was much higher than that of QVOA with a 300:1 ratio[77,153]. qPCR assays are unable to differentiate between defective and intact proviruses as well as integrated and unintegrated forms. Hence, they grossly overestimate the reservoir size.

To counter the inability of standard PCRs to differentiate between integrated and non-integrated forms, a different type of qPCR was developed which uses the Alu sequences found in the human genome to help detect integrated HIV-1 DNA. In Alu-PCR, one reaction is targeted towards the Alu regions in the human genome and the other reaction targets the GAG/LTR region to improve sensitivity[158-162]. The Alu-PCR is able to quantify the proviruses that are integrated[68,144,153]. As with qPCR, a standard curve is needed to quantify the results. This can be inefficient as not all Alu sequences will be in close proximity to the HIV-1 genome, therefore causing errors in detection, and the need for a correction factor to be calculated to account for this issue[155,158,160,161,163,164]..

Droplet digital PCR (ddPCR)

Droplet digital PCRs are more sensitive than qPCRs and give absolute quantification of the reservoir; therefore they were developed as an alternative method to measure the reservoir[145]. ddPCR involves normalizing DNA to a

particular concentration and adding this mix to special droplet generator oil that splits the liquid into nano-sized droplets such that each droplet (with or without a provirus) is its own PCR reaction[145]. Therefore, each provirus gets amplified, giving a more precise quantitative signal. ddPCR assays such as the GAG-LTR require a small sample size, are cost-effective, have a short turnaround time, are high-throughput, are more precise than qPCRs, are more tolerant to primer-probe mismatches than qPCR, and can be used to test samples from various tissues[64,65,68,144,145,151,155,165].

However, the single-plex ddPCR does have issues that make it difficult to use this assay in cure studies: Single-plex ddPCR assays are unable to differentiate between defective and intact proviruses as well as integrated and unintegrated genomes, and thus grossly overestimate the size of the reservoir. In some cases, the no-template control shows false positive results. Based on the assay design, it is not possible to sequence any of the positive droplets, as the reader discards all the material. Owing to the use of primers and the heterogeneity of the HIV-1 genome, primer-probe mismatches are an issue, but they are better tolerated with the ddPCR vs the qPCR. However, based on the demerits it is stated that the qPCR is a better assay to measure the reservoir studies vs the ddPCR [64,65,68,144,151,153,155,165].

Intact Proviral DNA Assay (multiplex ddPCR)

The intact proviral DNA assay is a multi-plex ddPCR assay designed to study the latent reservoir. In the IPDA, two regions of the HIV-1 genome are targeted: the packing sequence (psi, Ψ) upstream of the gag gene and the Rev Response

Element (RRE) in the Env gene (Figure 3). These two regions were chosen based on near full-length genome sequence (nFGS) analysis and bioinformatics; it was shown that any deletions in these regions indicate a high probability that the virus is defective[69]. The IPDA also includes a double quencher probe for hypermutations near the env region (most variable region[6]); if there is a hypermutation present, that provirus will amplify but not fluoresce, excluding it from the positive droplets[69].

The results for the IPDA are displayed in a 2D amplitude plot where three quadrants represent a type of provirus: quadrant 1 - gag intact only + non hypermutated; quadrant 2: fully intact (gag + env); and quadrant 4 - env intact only. Quadrant 3 consists of proviruses that did not amplify with the IPDA primers[69].

This assay also accounts for the RNase P30 gene (RPP30), whose PCR reaction determines the number of cells analyzed and the shearing index.

The primers for RPP30 bind at a distance of about 7kb from each other which is equivalent to the distance between the psi and RRE region. If the DNA gets fragmented during the processing for IPDA, it will be readily evident in the RPP30 PCR reaction by the presence of droplets in quadrants 1 and 4[69]. Based on nFGS, one study found that only 2.4% of the reservoir is truly intact (i.e., lacking deletions, hypermutations, insertions.) while 97.6% is defective[70,71]. The IPDA can detect and exclude the 97% of the proviruses that are deemed defective by nFGS which is much higher in comparison to standard GAG-LTR PCR (30%) and the Alu PCR (tends to zero). As verified by nFGS, 70% of the proviruses that are deemed intact by the IPDA are truly intact compared to 10%

detected by the Gag-LTR PCR. Some of the main advantages of the IPDA are: Its exclusion criteria as mentioned earlier[57,62,69,166]; its ability to detect increase in infected cell frequency as a result of clonal expansion[64,69,111,151,167]; and the strong correlation with the data generated by the QVOA with respect to half-life of the reservoir and the changes in frequencies of intact proviruses. The IPDA can be applied to frozen cell pellets[64 69,168]. It is cost-effective, requires small sample sizes and has a quick turnaround time[64,69].

However, the IPDA also has a few disadvantages: It is currently optimized for HIV-1 subtype B only[151,165,169,170]. It cannot differentiate between replication-competent and non-inducible intact proviruses since there could be defects in regions not overlapping the amplicons and the provirus could be integrated in an inactive gene[69,151]. It also overestimates (by ≈1.5 fold) the size of the reservoir for the same reason mentioned earlier. It also cannot differentiate between integrated and non-integrated forms of HIV-1 DNA, although the unintegrated forms are rare in patients on long-term effective ART, which adds to the overestimation of the size of the reservoir [64,69,70,151,169]. Primer-probe mismatches are possible due to single base mutations. However, since the IPDA is a multi-plex assay, primer-probe mismatches are readily identifiable vs standard PCRs[62,69,169,171,172].

Regardless, the advantages do outweigh the disadvantages and the IPDA is currently one of the best high-throughput PCR-based methods for differentiating between the intact and defective proviruses.

Quadraplex quantitative PCR

The quadraplex quantitative PCR assay also known as Q4PCR, was developed by Gaebler et al., using hundreds of proviral sequences from the Los Alamos Database to find the primers (in-silico). It consists of a limiting dilution of proviruses on which a long-distance PCR is performed paired with a multi-plex qPCR reaction where four regions of the HIV-1 genome: env, pol, psi (Ψ) and gag are interrogated, and the primers overlap the conserved portion of these regions[169,173]. The near full-length genome sequencing is only done on samples that show a positive reaction for two or more regions of the HIV-1 genome. This approach of sequence verification increases the probability of the assay to detect truly intact proviruses by eliminating the proviruses that are classified as intact in the qPCR but may have defects in the regions not overlapping the primers[64,165,169,173]. The Q4PCR is therefore a sensitive assay that can differentiate between intact and defective proviruses, and potentially can differentiate between integrated and unintegrated forms making it more specific than the IPDA[64,165,169,173]. For the Q4PCR, the cell input is detected by quantifying the amount of DNA added to the assay[169,173]. The Q4PCR however, is not a high-throughput assay as it involves near full-length genome sequencing, which is quite laborious.

The limiting dilution followed by long distance PCR makes the turnaround time relatively longer than the IPDA. Since the Q4PCR is based on the use of primer-probes, it is also subject to primer-probe mismatch issues, and was developed for subtype B. The Q4PCR can give insight into the integration site of the genome but cannot predict the inducibility of the provirus[64,169,173].

Sequencing methods

Sequencing approaches are extremely helpful in characterizing the reservoir of HIV-1 since they can help differentiate between defective and intact proviruses. There are several different techniques that are used to sequence the reservoir:

- Near full-length individual proviral sequencing (FLIP-seq): The provirus is amplified using an outer PCR followed by nested PCRs in segments. Sanger sequencing or next generation sequencing (NGS) is performed. The FLIP-seq can determine whether a provirus is genetically intact or defective, the defects contained by the provirus, and the contribution of clonal expansion to reservoir maintainence[64,86,103,151,174-176].
- MIP-seq/FLIP seq: These are newer techniques that involve amplifying the proviruses using multiple displacement amplification (MDA) followed by NGS and integration site analysis (MIP-seq). MIP-seq provides information about the intactness of the provirus as well as chromosomal integration site[64,103,151,177]

The disadvantages of sequencing techniques are that they are not high throughput; they are labor intensive and expensive; and in some cases, custom primers may be required[64,86,151,175]. However, sequencing techniques are able to characterize the reservoir well and can be used to supplement assays such as IPDA to characterize the reservoir. Gaebler et al., have merged a PCR-based method of measuring the reservoir to sequencing to help determine the replication competency of the proviruses in addition to their intactness and the size of the reservoir[169,173].

- *Assays that measure the concentration of transcriptionally/translationally competent viruses*

Tat/Rev Induced Limiting Dilution Assay

The *Tat/Rev* induced limiting dilution assay (TILDA) was developed to measure multiply spliced viral transcripts (transcriptionally competent) that can be produced by the latent reservoir upon stimulation using mitogens such as PMA and ionomycin [178]. In the TILDA, infected cells are stimulated with mitogens, and different dilutions of these activated cells are prepared followed by performance of a nested PCR for *Tat/Rev* regions[178]. It has been shown that Tat/Rev deletions are common in defective proviruses, which implies that several defective proviruses may not amplify these regions[73,178]. This approach of performing a PCR after culture permits skipping the RNA extraction step[178]. On comparing the frequency of latently infected CD4 T cells between TILDA, QVOA and PCR assays it was noted that the TILDA gave 48-fold higher values than QVOA but 6-27-fold lower values than PCR assays[65,178]. TILDA is a relatively simple assay compared to the QVOA as it requires less than a million cells, has a short turnaround time of 2 days, and RNA extraction is not needed[64,155,178]. It is sensitive, reproducible, and specific to HIV-1[178].

The TILDA, like the QVOA, depends on one round of stimulation, which may lead to underestimation of the reservoir[144,155,178,179]. However, since defective proviruses with no defects in the tat-rev regions also produce proteins and therefore viral transcripts, the TILDA is unable to differentiate the two and may overestimate the size of the reservoir[64,71-73,156,178].

This may be accounted for by using spontaneous production of transcripts by adding Tat to the cells rather than mitogens[178,180-182]. The TILDA is also more labor- and cost-intensive than PCR assays; however, it is more discriminating and can be used to study the effects of cure strategies on the reservoir[64,65,178].

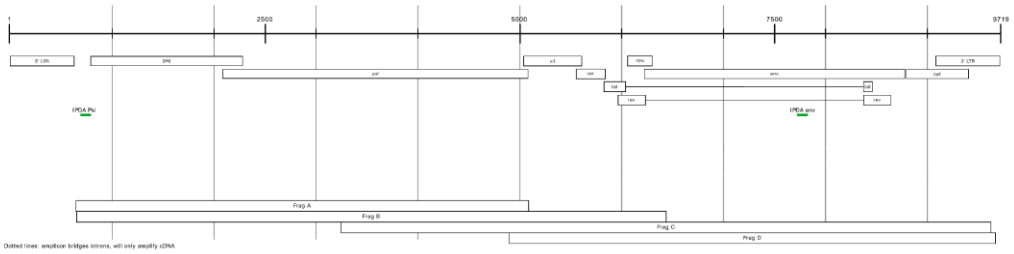


Figure 3. IPDA primer binding sites

The IPDA primers (in green) bind in the packaging signal region of the GAG gene and the Rev response element in the Env gene[69]. Created with Inkscape by Dr. Adit Dhummakupt.

Reservoir studies in Adult HIV-1 Infections using the IPDA.

Prior to the development of the IPDA, in studies done on adults who initiated ART during acute infection, across different cohorts, the following observations were made about the benefits of early effective ART vs treatment during chronic infection: reductions in the size of the HIV-1 reservoir[56,58,78,183-187], rapid immune reconstitution[58] and lower immune activation of T cells[56]. All of these studies were done using different assays[54]. Some studies showed that the reservoir decayed faster in adults who started ART early during the infection[61,184].

The development of the IPDA provided a high-throughput method to elucidate the dynamics of the different HIV-1 proviral species which would be helpful in informing cure strategies.

In a recent study where IPDA was applied to 81 participants who started ART during chronic infection[62], the authors noted differences in the decay rates between the two proviral species: intact and defective HIV-1 DNA. Over a course of the study, they observed that the intact proviruses decayed much faster than the defective HIV-1 proviruses, with a rate of -15.7 % per year for intact proviruses, in the first 7 years of virologic suppression (VS). In contrast, some participants demonstrated an expansion in 3' deleted/hypermutated HIV-1 DNA and some showed an expansion in the 5' deleted HIV-1 DNA population during the first seven years after VS. After the seven-year inflexion point, the rates of decay decreased for all species, but were still higher for intact proviruses compared to defective HIV-1 proviruses. The main conclusion was that the decay rates differed between the proviral species with intact HIV-1 DNA decaying much faster than the defective HIV-1 species especially in the first 7 years after VS[62].

This suggests that the intact replication competent proviruses stochastically reactivate, express viral proteins and are killed either by the HIV specific immune responses or by viral cytopathic effects[62]. However, the infection of new cells by the virions produced by this reactivation is prevented by ART.

In a second study[57] on 44 participants on effective ART over a course of 12 years after ART, samples were collected at a median of 7.1 years after ART initiation (time point 1), 3.7 years after time point 1 and 1.8 years after time point

2. It was observed that although the intact proviral reservoir declined, the defective HIV-1 proviral reservoir either expanded, remained stable or declined. Owing to this, the total HIV-1 DNA remained relatively stable over the course of effective ART, even though the intact proviral reservoir seemingly declined. They also observed that the median intact copies/million CD4 T cells decreased from time point 1 (median 7.1 years; T1) to time point 3 (median 12.6 years; T3). The %intact fell from 9.8% at T1, to 5.8% at T3 and, some of the participants showed an increase in the 3' deleted/hypermutated or 5' deleted HIV- 1 proviruses while some showed a decline, between T1-T3.

These data support the idea that the replication-competent proviruses stochastically reactivate, express proteins and are then eliminated by the HIV specific immune responses or are killed as a result of viral cytopathic effects. Defective proviruses do produce proteins and express them on the surface of the cells; this is observed with 5' deleted proviruses which can decline slowly in some cases, but their loss can be compensated by expansion or stability of 3' deleted/hypermutated proviruses. This also supports the idea that the defective proviruses expand over time on suppressive ART allowing the total HIV-1 DNA concentration to remain stable over time[167,188].

Altogether this study showed that while the intact proviruses decay over time on ART, the defective proviruses continue to expand and dominate the reservoir landscape. Over the course of the study, median change per year for the 3' deleted/hypermutated HIV-1 proviruses, was found to be zero. This proved that little to no net change was observed for this set of proviruses[57].

A few interesting observations in the study were:

1) Low CD4+ T cell counts prior to ART initiation were found to be associated with higher total HIV-1 DNA copies and a higher proportion of defective HIV-1 DNA copies on effective ART. The authors speculated that this was a result of longer untreated infection which permitted the virus to accumulate over time with predominantly defective HIV-1 proviral species.
2) Lack of decline of defective HIV-1 proviruses in the face of declining intact was speculated to be caused by multiple factors such as integration of intact proviruses into transcriptionally active sites permitting expression of viral proteins/virion release and clearance by the immune system/viral cytopathic effects. Expansion of cells carrying the defective HIV-1 proviruses could dilute the intact HIV-1 proviral reservoir (being eliminated) making it appear as though the defective HIV-1 proviruses remain stable over time. HIV-1 is known to integrate into genes involved in cell cycle control which could explain why proliferation may occur[111,167].
3) The half-life of the intact proviral reservoir detected by IPDA (7.1 years) was much longer than that determined by the QVOA (3.6-3.7 years), which could indicate that a portion of the intact proviruses have deep latency features [50,90]. This has implications in terms of clearance of the reservoir.

Will these proviruses be reactivated after multiple rounds of stimulation?

The answer to the above question was found in another study done on 12 adults with HIV-1, on long-term suppressive ART, who maintained plasma viral load<20 copies/mL on ART for >6 months[116]. In this study, Hosmane et al.,

used a modified version of the QVOA where instead of just one round of stimulation, four rounds of stimulation were done.

They observed that during the first round of stimulation only 60% of the culture were positive for p24 even though >99% of the resting CD4 T cells had proliferated which would indicate that cells without replication-competent proviruses might also have proliferated. However, each additional round of stimulation was beneficial in reactivating more viruses that would not have been reactivated without the additional stimulation[116]. This is an important finding as it proves that the latent reservoir is capable of proliferating without reactivation on initial exposure to mitogens but upon additional stimulus, it reactivates to produce virions. The results of this study support the following ideas: 1) the reservoir size lies in between that measured by assays such as IPDA and QVOA, and more than one assay will be required to quantify the effect of cure strategies on the reservoir; and 2) the curative strategies will need to account for the lack of viral gene expression by a proportion of the replication-competent proviruses in spite of proliferation.

Since the IPDA is a relatively new assay, Simonetti et al., sought to determine the performance characteristics of the IPDA by applying it to 400 adults on ART from different cohort studies[171]. The main observations were as follows:

1. Of the three DNA species, the proviral load was the lowest for the intact proviruses and only 8% of the total HIV-1 DNA was composed of intact proviruses similar to what was seen in the study by Gandhi et al.,[171,189]. These observations confirmed that on ART, the proviral landscape is dominated by cells carrying defective proviruses while the frequencies of

cells carrying intact proviruses are relatively low. The authors found the ratio of defective to intact proviruses to be ~12.5:1[171].

2. The frequencies of intact proviruses reported by the IPDA are 50-fold greater than that observed in QVOA and 12.5-fold lower than the total frequency of infected cells.

   Therefore, the IPDA gives a better estimate of the reservoir size than qPCRs (overestimation) and QVOA (underestimation). However, some of the proviruses labelled as intact by the IPDA may have minor defects in regions not overlapping with the amplicons of the IPDA[171] which could affect their ability to reactivate and produce infectious virus.

3. Since the IPDA uses primer-probes for detection, it is susceptible to sequence polymorphisms in the regions that overlap the primer-probe binding sites. Bruner et al., had shown that 4% of proviruses are likely to not be amplified due to defects in the regions overlapping both the primer-probe sets i.e., gag (psi) and env[69]. This meant that only 96% of the proviruses would be detected by the IPDA. Of these 96%, Simonetti et al., observed that a total of 6.3% of participants (25/400), out of 400, showed amplicon signal failure with 3.5% (14/400) showing amplicon signal failure for psi and 2.8% (11/400) showing amplicon signal failure. However, since the IPDA is a multi-plex PCR assay, it was easy to catch the signal failure relative to single-plex PCR assays. This was resolved by developing individual primer-probe sets for the participants that showed signal failure. This was an important observation as it showed that sequence polymorphisms are common and expected in studies involving large cohorts and identifying

common polymorphisms could allow us to develop primer-probes that are less susceptible to signal failure[171].

In another study, Gaebler et al., sought to compare the IPDA to an assay developed in their laboratory known as the quantitative quadraplex PCR assay (Q4PCR)[169,173]. They applied the IPDA and Q4PCR to 39 participants on long-term suppressive ART (median of 8.4 years). They observed the following:

1. The frequency of cells containing intact proviruses detected by Q4PCR (median of 5 copies/million CD4 T cells) was between that detected by IPDA (median of 65 copies/million CD4 T cells) and QVOA (median of 0.6 infectious units/million CD4 T cells)[169].
2. A positive correlation between the two assays was observed with the caveat that IPDA detects almost 19-fold higher frequencies of intact proviruses than Q4PCR does. This was attributed to: a) the inefficiency of the initial long-range PCR amplification; b) the misclassification of proviruses that may have defects in regions not overlapping the IPDA primer-probes as intact by the IPDA; and c) cell normalization methods for both the assays[169].
3. Using Q4PCR, they were able to show the high specificity of psi and env primer-probe combinations (same as that used in IPDA) for excluding defective proviruses. Of the 3 participants that were shown to belong to another subtype by the Q4PCR, the IPDA showed primer-probe mismatches in the env region reenforcing the idea that IPDA, which was developed for subtype B, must be optimized for other subtypes before use[169].

4. Amplicon signal failure was observed in 18% of the participants (7/39) for the IPDA, while the Q4PCR was unable to retrieve intact sequences from 20.5% of the participants (8/39).

The main conclusion was that the differences in the two assays continue to support the idea that the proviral landscape is dominated by defective proviruses with low frequencies of intact proviruses and using the two assays together would help provide a deeper characterization of the reservoir[169].

In a study by Falcinelli et al., the authors sought to compare the IPDA to the QVOA. They tested resting CD4 T cells from 83 participants on suppressive ART in a cross-sectional and longitudinal study. Of the 83 participants, 7 (8.4%) showed primer-probe mismatches. Of the other 76, 16 were treated during acute infection and 60 were treated during the chronic infection.

The cross-sectional study revealed the following:

1. As expected, the frequencies of proviruses were the lowest for QVOA, followed by intact, 5' deleted, 3' deleted/hypermutated and total HIV-1 DNA. A median of 11.6% of total proviruses consisted of intact proviruses.
2. In the early-treated participants, the proportion of intact DNA in the HIV-1 reservoir was much higher than that in chronically infected participants. This could indicate that the starting levels of intact HIV-1 DNA in early-treated participants are higher. Another possibility was that the IPDA was developed for detection of proviruses during treatment of chronic infection. There are differences in hypermutations and deletions between early treated and late treated participants. These differences could have affected the ability of the IPDA to detect intact proviruses in acute infections, to confirm this idea, more research would be required.

3. When a correlation with QVOA was run, it was found that intact and total HIV-1 DNA had a better association with replication-competent virus outgrowth than with defective proviruses. This was to be expected since the IPDA is capable of detecting over 96% of the intact proviruses.

In the longitudinal analysis, the following observations were made:

1. In early treated participants, the frequency of decrease for intact proviruses detected by IPDA corresponded to that of the replication-competent proviruses detected by QVOA with increasing durations of suppression. The frequencies of defective proviruses were variable and showed either expansion, decline or stability with increasing duration of suppression. This observation supported the idea that defective proviruses dominate the proviral landscape during long-term treatment.
2. In participants who were treated during chronic infections, the intact proviruses as quantified by the IPDA showed two trends. In most cases, the intact proviruses declined in parallel to the replication-competent proviruses; however, in a few cases, the intact proviruses remained stable as the replication-competent proviruses declined. The reasons for this could be variance in the assay or clonal expansion of intact proviruses[176,190]. Clonal expansion of intact proviruses without reactivation of the replication-competent proviruses is possible, as discussed above, and could contribute to the persistence of the replication-competent reservoir[116].

The main conclusions of the study were: intact proviruses are selectively eliminated, the intact proviral frequencies detected by the IPDA correlate well with those detected by QVOA; the percent decrease in intact proviruses as

detected by the IPDA parallels percent decrease of replication-competent proviruses seen with QVOA; and in comparison, to the other assays, the results of the IPDA provide a better upper limit to the size of the reservoir.

In conclusion, the IPDA has proven to be a beneficial high-throughput tool to quantify and characterize the dynamics of the reservoir in adults.

Reservoir studies in Pediatric HIV-1 Infections

The latent reservoir in pediatric infections was discovered in 2000 using the QVOA[74]. Since then, many studies have been performed to characterize the reservoir and the effect of ART on the reservoir[36,55,59,60,63,94,179,191-220].

Initially the studies describing the decay dynamics of the reservoir in perinatal infections were performed on samples from participants who were treated during chronic infection (>2years after infection)[208,211]. Initiation of ART in children lead to a biphasic decay where the rate of decay was rapid prior to the inflexion point, 4 weeks after ART initiation for study by De Rossi et al., after which the rate of decay slowed down[218].

In 2005, a clinical trial namely Children with HIV early antiretroviral therapy (CHER) was started[203]. In this trial, children aged 6 -12 weeks (median age = 7.4 weeks) were divided into two arms: one arm deferred therapy and one arm received time limited ART for either 40 or 96 weeks and then ART was interrupted.

After a median follow up of 40 weeks, it was observed that: 1) mortality rates decreased by 76% for early treated vs deferred treatment; 2) the CD4 T cells

were reconstituted in the early treated group vs the deferred treatment; and 3) the disease progression declined by 75% for early treated participants vs deferred treatment[203].

The results of this seminal study were supported by other observational studies where it was found that early ART: 1) reduced the chances of disease progression[191,200,214]; 2) The proviral loads for early treated participants were lower as compared to those who received deferred treatment[191]; and 3) the levels of CD4 T cells were higher in early treated participants compared to those who received deferred treatment[191]. This set the stage for changing recommendations for timing of ART initiation set by the WHO[221].

Following these new guidelines, the field moved towards early treatment and in some cases very early treatment (e.g., Mississippi baby was treated within 30 hours of birth[92]). The data from these studies, across several cohorts, subtypes and durations of suppression supported the following observations: 1) the reservoir size in early treated participants was smaller than that observed in late treated participants[216,220,222]; 2) the proviral loads were very low and, in many cases, undetectable within months of virologic suppression for the early treated participants compared to the late treated[55,59,191,196,199,201,202,207]; 3) early treatment led to early suppression of viremia, thus reducing the size of the reservoir and immune activation caused by HIV-1 replication and antigen exposure[199,201,219]; and 4) disease progression was curtailed[215].

However, the reason for decrease in the size of the reservoir was not well elucidated therefore, to understand what was going on, in a seminal study, Uprety et al., studied the decay dynamics of the reservoir in long-term

suppressed children. In their study, Uprety et al., looked at two groups stratified by age at virologic suppression, in the Pediatric HIV AIDS Cohort Study[60]. Both groups had participants that were early treated but one group consisted of participants that were suppressed before one year of age and the other group consisted of participants suppressed between 1-5 years of age. It was observed that at the time of suppression, both the groups had similar amounts of HIV-1 DNA as detected by the Pol-LTR ddPCR assay. However, within the first two years of virologic suppression (VS), the HIV-1 DNA decayed much faster in early suppressed group than it did in the later suppressed group[60]. This explained the reason for the eventual smaller reservoir sizes in early treated children.

As the study used a total HIV-1 DNA measure, the reason for the faster decay of HIV-1 DNA in the early suppressed group was not identified[60]. This is where the IPDA assay could bridge the gap, by allowing characterization of the HIV-1 DNA species and providing a rationale for the faster decline of HIV-1 DNA in the first 2 years after VS with early treatment[60,69]. The studies done on the reservoirs in perinatal HIV-1 infections have applied assays such as QVOA[55,63,193,198], near full-length genome sequencing[193,195,219],qPCR or ddPCR of the gag/pol/LTR[59,60,63,193,196,199,202,205,215,216,219,220].

QVOA as mentioned earlier gives the minimum estimate of the size of the reservoir and requires large volumes of blood, which makes it difficult to study the reservoir in very early treated pediatric participants. It is also not feasible for large cohort studies[64,65,68,144,148,151,153,155]. The near-full length genome sequence (nFGS) on the other hand, has shown that there is a paucity of intact

proviruses in early treated children[193,195]. However, the long-distance PCR of the nFGS is inefficient and could preferentially amplify shorter strands of defective proviruses, leading to an underestimation of the intact proviruses[64,86,151,169,173,175]. Therefore, nFGS is not well suited to accurately quantify and characterize the reservoir in children.

Total HIV-1 DNA measures such as with qPCR or ddPCR target single-amplicon regions such as the gag, pol, LTR regions [59,60,64,65,68,144,145,151,153,155,165,205]. These assays can be particularly useful in measuring the reservoir size during long-term suppressive therapy when the proviral loads fall to exceedingly low levels, below 5 HIV-1 DNA copies/million PBMCs and for which assay optimization for detection across all subtypes are critical[59,60,63,196,205]. However, total HIV-1 DNA assays overestimate the size of the reservoir as they also quantify defective proviruses which are known to accumulate as the duration of suppression increases[64,65,68,144,145,151,153,155,165]. Feasibility of measuring the proviral reservoir in children is important; however, we need to improve the understanding of the mechanism behind differential decay dynamics of the reservoir in early vs late treated children.

The application of the IPDA to perinatal infections will help us to understand the cause behind differential decay dynamics in early and late treated participants by characterizing the reservoir. What leads to the differential decay of the reservoir in early vs late treated especially when they start out with the same proviral loads at the time of suppression?[60] Is it the clearance of the intact proviruses? Or does clonal expansion of the defective proviruses dilute the

reservoir?[70,167] The questions can be answered by the IPDA and will inform the development of remission strategies.

To our knowledge, no study has been done on pediatric HIV-1 infections using the IPDA, and the present study is the first preliminary study that will allow us to understand the dynamics of the different species of HIV-1 in the reservoir as a function of age at virologic suppression and duration of virologic suppression.

Study Aims:

1. Optimize and validate the Intact proviral DNA assay for Pediatric samples where cell availability for study is limited.

2. Estimate the proviral reservoir size, and the distribution of total, intact, and 3' deleted/hypermutated and 5' deleted HIV-1 proviral genomes during effective ART in children and adolescents living with perinatally acquired HIV-1.

3. Compare the proviral reservoir size of total, intact, and 3' deleted/hypermutated and 5' deleted HIV-1 proviral genomes as a function of age at virologic suppression and duration of virologic control in children and adolescents living with perinatal HIV-1.

II. Materials and Methods

1. Establishing an Internal Standard for the IPDA

The JLAT cell line was obtained from the NIH AIDS reagent program. JLAT cells also known as Jurkat cells are a cell line derived from a person with T cell leukemia [223]. A derivative of the JLAT cell line was developed by Jordan et al., to study the HIV-1 reservoir in an in-vitro model[224]. An HIV-1 genome was inserted into the cell line and was tagged with green fluorescent protein. The JLAT cells do not produce infectious virus owing to a frameshift mutation in the env region[224]. However, this cell line is ideal for IPDA since each cell only carries one copy of the HIV-1 genome, permitting its use for quantitative studies[224,225]. The methods described below for maintaining JLAT cells culture were optimized in the laboratory[69].

1.1. Preparation of JLATS for testing

Thawed JLAT cells were spun at 300 x g for 10 min using the Beckman Coulter Allegra X14R, the supernatant was decanted, and the cell pellet was resuspended in 10mL of Thawing media (TM; 50% FBS (Sigma Aldrich; #19G462) + 50% RPMI 1640 (Gibco; #61870-036)). The cells were spun again at 300 x g for 10 min, the supernatant was decanted, and the cell pellet was resuspended in 10mL of JLAT media (RPMI 1640 + 10% heat inactivated Fetal Bovine Serum (FBS) + 1% Pen Strep (Gibco; #15140-122)). The cell count and viability check were then performed using a 1:2 dilution with trypan blue (Gibco; 15250-061).

10µL of this trypan blue:cells mix was added to the hemocytometer (Bulldog Bio; #DHC N005). The cell count was determined by counting the number of JLAT cells in the four WBC quadrants of the hemocytometer slide. After this, cells/mL had to be determined as well as the viability of the cells. The formula used to determine cell/mL was as follows:

$\frac{Cells}{mL} = \frac{C}{4} \times 10{,}000 \times dilution\ factor$, where C = Number of cells, 4 = number of quadrants and 10,000 represents the volume of each quadrant in mL

$$Viability = \left(\frac{Number\ of\ live\ cells}{Total\ Cells}\right) * 100$$

Based on the cell count, the cells were spun at 300 x g for 10 min and then resuspended in the appropriate amount of JLAT media to achieve a concentration of 1-3*10^5 cells/mL. The cells were incubated in the 37°C incubator, 5% CO_2, humidified air, for 48 hours.

1.2. Passaging of JLATS

The cell density and viability of the JLATs were checked every 2-3 days using trypan-blue exclusion dye staining. The cell count and viability check were performed using a 1:2 dilution of JLATs to trypan blue as described above. If the slide was too crowded (i.e., more than 100 cells in one WBC quadrant of the hemocytometer), a 1:10 dilution of cells to trypan blue was created. The cell count was performed and the viability of the JLAT cells was determined as described above. We determined that the optimal cell concentration had to be between 2*10^5 cells/mL and 2*10^6 cell/mL to ensure that the JLATs did not start losing viability.

If the cell count was close to 5 million cells/mL, all the cells were passaged and split into the appropriate number of flasks. The optimal concentration of the cells after passaging and splitting was determined to be 1-3*10^5 cells/mL. The flasks were incubated in the 37^0C incubator, 5% CO_2 for 48 hours.

1.3. Maintaining frozen stocks of JLATS

The cell density and viability of the JLATs were checked every 2-3 days as described above. When the cell density was approximately 5*10^6 cells/mL frozen stocks of the JLAT cells were created by spinning down the cells, adding freezer medium (FBS: 90% + DMSO:10%) to the cells. The cell suspension was then dispensed in 1mL aliquots into cryovials which were placed into Mr. Frosty (slow cooling container filled with isopropyl alcohol to achieve a cooling rate of 1^oC/min; Nalgene; #5100-0001) and then kept in the -80^0C freezer overnight. The following day the vials were placed in the vapor phase nitrogen tank for long-term storage.

2. Preparation of JLAT Standards

For preparing the JLAT standards, a fresh blood bag was ordered from the New York Blood Centre on the day of preparation to have stocks (pellets of 5 million cells) of human peripheral blood mononuclear cells (PBMCs) to spike with known concentrations of JLAT cells. The blood from the blood bag was overlayed on Ficoll and spun at 400 x g for 25 min using the Beckman Coulter Allegra X14R. After the blood was divided into the four layers, the buffy layer containing the PBMCs was extracted. The PBMCs were then washed (wash

media: phosphate buffered saline (PBS), pH7.4, 2% heat inactivated new-born calf serum, 0.1% glucose, 20U/mL penicillin, 20µg/mL streptomycin, 12mM HEPES, pH7.4) and the cell count was performed as described above.

Once the cell count was determined, wash media was added to the PBMCs to achieve a concentration of 10 million PBMCs/mL which was confirmed with another cell count.

Based on the final count, the number of aliquots for JLAT standards that could be prepared was determined. 4mL of the PBMCs was aliquoted out and split into tubes containing 0.3mL of the PBMCs (3 million cells). The tubes were kept aside. This was the I9 standard (no HIV-1 DNA standard).

Next, the JLATs were spun at 300 x g for 10 min, and the supernatant was decanted. The cell pellet was resuspended in wash media to achieve a concentration of 1 million JLATs/mL, this was verified by another cell count. To begin generating the standards, 4 mL of the 10 million PBMCs/mL stock (corresponding to 40 million PBMCs) was taken and added to a 15mL conical tube. To this 0.04mL of the 1 million JLAT cells/mL stock (corresponding to 40,000 JLAT cells) was added to give a solution containing 1000 JLAT cells/million PBMCs. The remaining standards were prepared using the table below:

Dilution	I1 (1000)*	I2 (100)	I3 (25)	I4 (10)	I6 (5)	I7 (2.5)	I8 (1.25)	I9 (0)
Initial volume of PBMCS (mL)	4	3.6	3.1	2.4	2	2	2	4
Volume of previous dilution added to the tube (mL)	-	0.4	0.9	1.6	2	2	2	-
Dilution Factor	10	4	2.5	1.5	2	2	2	0
Total Volume (mL)	4	4	4	4	4	4	4	4
Volume remaining after serial dilution(mL)	3.6	3.1	2.4	1.6	2	2	4	4
Expected Copies of HIV-1 DNA/million PBMCS	1000	100	25	10	5	2.5	1.25	0
* To I1, 0.04mL of 1 million JLATs/mL stock was added.								

Table 1: JLAT Standard Preparation

The dilution factors in the table refer to the dilution that was made from the respective columns for the next concentration. In column 1, 10 was the dilution we used to make I2.

Each sample was mixed well to ensure even JLAT distribution. Once the dilutions were ready, the cells for each dilution were aliquoted out as pellets of either 5 million or 3 million cells (PBMCs+ JLATs). The tubes were then spun at 3000 rpm for 5 min using the tabletop Eppendorf Centrifuge 5430R and most of the supernatant was removed using a pipette, leaving behind about 50µL. The tubes were stored at -80^{0}C in a box.

3. Isolation of Genomic DNA from JLAT/PBMC Standards

The Qiagen Qi-AMP Blood DNA Midi Kit (#51183; Qiagen, Valencia, California) was used to isolate genomic DNA from the JLAT standards.

Once the tubes containing the pellets were labelled, 1.2mL of Buffer AL (Lysis Buffer; provided in the kit) was added to each cell pellet to reduce the chances of breakdown of DNA by DNAses or shearing. 100µL reconstituted Proteinase K (provided in the kit) was pipetted out into each of the sterile 15mL centrifuge Tube (VWR). The contents of the original cell tubes were transferred into the respective 15mL centrifuge tubes. The cell tubes were then rinsed with 1mL 1X PBS (Quality Biological; #114-058-101) and the original cell tubes were discarded. The 15mL centrifuge tubes were then pulse vortexed using the vortex mixer (Fisher Scientific Vortex Mixer) and were placed upright in the 70^{o}C Isotemp Waterbath (Fisher Scientific Isotemp Water bath) for 10 min. After the incubation, the tubes were spun at 3200 rpm for 3 min using the Jouan C3i tabletop centrifuge. Post-centrifugation, 1.2mL of 100% ethanol (Fisher Bioreagents; #BP2818-500) at room temperature was added to each of the tubes. The tubes were then pulse vortexed and spun down at 3200 rpm for 3 min.

The contents of each standard's tube were transferred carefully into the respective Qiagen MIDI centrifuge tubes with spin column. The tubes with spin columns were spun at 3200 rpm for 5 min. For each standard, the spin columns were carefully taken out, and the flow through was discarded. The collection tubes were retained, and the spin columns were re-inserted into their respective tubes. 2mL of Buffer AW1(provided in the kit) was added into each spin column followed by a spin (4000 rpm for 4 min), and then 2mL of Buffer AW2 (provided

in the kit) was added into the spin columns. The tubes were spun at 4000 rpm for 18 min and then again for an additional 4 min. The spin columns were then taken out and placed into new 15 mL MIDI centrifuge tubes. Next, 200µL Buffer AE (provided in the kit) was added to the spin columns and the buffer was allowed to infiltrate the filter before centrifuging.

The new tubes with spin columns were spun at 4000 rpm for 4 min. This step was repeated 4 times for a total of 1mL eluate. 1µL of thawed glycogen (Thermofisher Scientific, #R0561) was added to two pre-labeled 1.5mL microcentrifuge tubes for each standard while all tubes were still placed on ice. Next, 500µL of eluted DNA was transferred to the respectively labelled two 1.5mL microcentrifuge tubes. 50µL 3M sodium acetate (Quality Biological; #351-035-721) and 1mL cold 100% Ethanol were added to each of the 1.5mL tubes containing eluted DNA. The tubes were gently inverted to promote mixing and were then stored overnight in a -20°C freezer.

The tubes were taken out from the freezer and spun in an Eppendorf Refrigerated Microcentrifuge 5417R at 4^0C, 14,000 rpm for 30 min. After the spin, the supernatant was decanted, and the DNA pellets for each of the samples respectively were combined into one tube using 70% ethanol. The tubes containing the pellets were centrifuged at 14,000 rpm (room-temperature) for 2 min and the supernatant was decanted. The tubes without pellets were rinsed with 1mL 70% Ethanol and the contents were added to the tubes with the pellets. The tubes containing the pellets were spun at 14,000 rpm (room-temperature) for 2 min and the supernatant was decanted. The tubes were spun again at 14,000 rpm (room-temperature) for 1 min, and any remaining supernatant was removed carefully.

The DNA pellet was left to air-dry. Once the alcohol had evaporated, the DNA pellet was resuspended in 50µL of Buffer AE.

The optical density was measured on the Thermo Scientific Nanodrop Lite Spectrophotometer and recorded. The tubes were then stored in the working box located in the -20^{0}C freezer [60,205].

The Qiagen Gentra Puregene Blood kit (Cat # 158467; Qiagen, Valencia, California) was used to isolate genomic DNA to reduce shearing.

The tubes containing the pellets were labeled and 0.6mL of cell lysis solution (provided in the kit) was added to each pellet to reduce the chances of breakdown of DNA by DNAses or shearing. The suspension was mixed by pipetting. To this, 3µL RNase A solution (provided in the kit) was added and the tubes were mixed by inversion 25 times. The tubes were incubated for 5 min at 37^{0}C followed by incubation on ice for 2 min. 200µL of Protein Precipitation Solution (provided in kit) was added and the tubes were vortexed vigorously and spun at 16000 x g for 3 min using the tabletop Eppendorf Centrifuge 5451D. 600µL of isopropanol was added to fresh 1.5mL microcentrifuge tubes and the supernatants from each of the tubes were transferred to the respective tubes containing isopropanol. The tubes were mixed by gently inverting them 50 times and were then spun at 16000 x g for 1 minute. The supernatant was discarded, and the remaining liquid was drained from the tube onto an absorbent paper. 300µL of ethanol was added to the tubes, the contents were mixed, the tubes were spun at 16000 x g for 1 min and the supernatant was discarded. The ethanol wash was repeated one more time. After the 2nd wash, the pellet was allowed to air dry for 5 min. Next 50µL of DNA hydration solution (provided in the kit) was added, the tubes were vortexed for 5 seconds and incubated at

65°C for 1 h using the Eppendorf Thermomixer 5350. This was followed by overnight incubation at room temperature with gentle shaking. The samples were then spun at 16000 x g for 1 min and stored in the working box located in the -20^{0}C freezer.

For each Pediatric HIV/AIDS Cohort Study (PHACS) sample being tested, four controls were required, two high concentration controls I1 and I2 with 1000 and 100 HIV-1 DNA copies/million PBMCs respectively, a negative control I9 with no HIV-1 DNA copies/million PBMCs and a no template control (NTC). Since the DNA from the PHACS samples was previously isolated, only the controls for the same were isolated using the JLAT standard.

4. Intact Proviral DNA Assay

The Intact Proviral DNA Assay (IPDA) is a multiplex droplet digital PCR (ddPCR) assay[69]. With ddPCR, every droplet is considered as an individual PCR (even though not every droplet has a proviral genome); therefore, each droplet is analyzed individually. This provides higher resolution as compared to the qPCR.[145] In the IPDA , the first reaction is targeted to two highly conserved genes in the HIV-1 genome called psi (packing signal) upstream of the gag gene and the rev response element (RRE) in env.[69] The second reaction is targeted to the housekeeping gene RNAse P30 (RPP30) for determining the number of cells analyzed as well as to provide estimates of DNA shearing[60,69,145]. The primer-probe sequences for RPP30 were based on previously published methods by Kinloch et al[172].

4.1 Preparing the Master Mix

Primer Probe Preparation

The primer and probe were reconstituted to a concentration of 100μM before use. Table 2 summarizes the primer-probe sequences used in the assay. The 20X primer-probe solutions derived from the 100μM solutions were used in the assay.

Each sample was run in 8 replicates to assess for gag (Psi), env and hypermutations; one replicate reaction was set up for RPP30; along with the controls (I9) and no template control (NTC); these were run in 4 replicates. In ddPCR, the data from droplets in every well for each sample is aggregated using Poisson distribution to give a single value.

The master mix was prepared by taking the desired volume (based on the number of replicates) of: ddPCR Supermix (Bio-Rad, #186-3024) + 20X primer probe solutions + DNase/RNase free water. This mix was then vortexed and spun briefly.

4.3 Preparation of the ddPCR Plate (Assay Master Mix + Samples)

The samples for the GAG/ENV/Hypermut PCR were prepared by mixing 135μL of the GAG/ENV/Hypermut master mix and 45μL of normalized DNA.

The controls for this PCR were prepared by mixing 22.5μL of the GAG/ENV/Hypermut master mix and 7.5μL of normalized DNA. The samples for the RPP30 PCR were prepared by mixing 22.5μL of the RPP30 master mix and 7.5μL of normalized DNA for samples and controls.

Assay	Primer/Probe Name	PCR Reaction	Primer/ Probe Type	Primer Sequence
Hypermut	IPDA Hypermut Probe			5'-/5IABkFQ/CCT-TAG-GTT-CTT-AGG-AGC/3IABkFQ/-3'
GAG	IPDA GAG Fw	Multiplex	Forward	5' – TCT-CGA-CGC-AGG-ACT-CG – 3'
	IPDA GAG Rv		Reverse	5' — TAC-TGA-CGC-TCT-CGC-ACC-3'
	IPDA GAG Probe		FAM	5'-/56-FAM/CTC-TCT-CCT/ZEN/TCT-AGC-CTC/3IABkFQ/ 3'
Env	IPDA Env Fw	Multiplex	Forward	5'- AGT-GGT-GCA-GAG-AGA-AAA-AAG-AGC-3'
	IPDA Env Rv		Reverse	5'- GTC-TGG-CCT-GTA-CCG-TCA-GC-3'
	IPDA Env Probe		HEX	5'-/5HEX/CCT-TGG-GTT/ZEN/CTT-GGG-AGC/3IABkFQ/
RPP30 Shear	IPDA RPP30 Shear Fw	Multiplex	Forward	5'- CCA-TTT-GCT-GCT-CCT-TGG-G-3'
	IPDA RPP30 Shear Rv		Reverse	5'- CAT-GCA-AAG-GAG-GAA-GCC-G-3'
	IPDA RPP30 Shear Probe		FAM	5'- /56-FAM/AAGGAGCAA/ZEN/GGTTCTATTGTAG/3IABkFQ/-3'
RPP30	IPDA RPP30 Fw	Multiplex	Forward	5'- GAT-TTG-GAC-CTG-CGA-GCG-3'
	IPDA RPP30 Rv		Reverse	5'- GCG-GCT-GTC-TCC-ACA-AGT-3'
	IPDA RPP30 Probe		HEX	5'-/5HEX/CTGACCTGA/ZEN/AGGCTCT/3IABkFQ/-3'

Table 2: Primer Probe Sequences for IPDA[172]

The table shows the different sequences used by the PCR reaction. The GAG and ENV probes fluoresce in different channels based on the presence of the amplicon within a given droplet. The hypermutation probe is a double quencher probe and therefore will show no fluoresce if hypermutations are present in the PCR amplicon[69].

4.4 Droplet Generation

The droplet generation was started with the “GAG/ENV/Hypermut” plate. The samples were mixed by pipetting up and down 3 times. Then 20µL of master mix and DNA mixture was transferred to middle row of wells of the cartridge for droplet generation. The wells were checked for bubbles. All bubbles were burst carefully using the micropipette tip. Next, the Bio-Rad droplet generation oil was pipetted up and down once to wet the pipette tip and 70µL of droplet generation oil (Bio-Rad; #186-3005), was dispensed from the solution basin (VWR; # 613-1175) to bottom wells of cartridge. If any wells were not used in a column, the empty wells were filled with 20µL of 1X buffer control (1:1 supermix and water). A generator gasket (Bio-Rad; #186-3009) was hooked on to cartridge holder. The QX200 droplet generator was opened and the adaptor containing the cartridge and gasket was placed into the unit for droplet generation.

Next, a multichannel pipette was positioned at ~30-45° angle into the topmost wells of the DG8 cartridge (droplet generator cartridge; Bio-Rad; #186-4008), which contained the freshly generated droplets. Approximately 44µL from the droplet wells was drawn into the pipette tips and then slowly dispensed into the appropriate column of 96-Well semi-skirted PCR plate based on plate layout on the “IPDA Analysis Template.” This process was repeated for all samples and controls in the “GAG/ENV/Hypermut” and RPP30” plates as needed. Once the plate was fully loaded, the plate containing generated droplets was placed onto the sealing stand of the PX1 PCR plate sealer instrument. [59,60,145].

4.5 PCR reaction and Plate Reader

The sealed plate was placed in the 96-well Bio-Rad C1000 Touch thermocycler. The thermocycler was set for 50% ramping speed (2°C/sec) and 105^0C heated lid. The chemistry of droplets requires a lower ramping speed to achieve the temperature required for the PCR properly[172] and the following program was run:

Temperature	Duration	
95°C	10min	
94°C	30sec	60 Cycles
53°C	60sec	
98°C	10min	
4°C	∞	

The plate sat in the thermal cycler overnight. Following the completion of PCR, the template on the Quantasoft software of the reader was prepared using the following settings:

Setting	Label
RED	Rare detection event
ddPCR Super-Mix (no dUTP)	
Channel 1 FAM	GAG/RPP30 Shear Unknown
Channel 2 HEX	ENV/RPP30/Unknown

Post template preparation in the software, the plate was read.

4.6 Analysis

The results of the IPDA were analyzed using the Quantasoft Software version 1.7.4.0917 by Bio-Rad Industries. The first step involved determining the number of droplets generated in each reaction well. The baseline level accepted was 6000 droplets which was determined empirically from prior experiments done in the laboratory. Any well that had less than 6000 droplets was labelled as “EXCLUDE” to eliminate it from analysis as this indicated that there was a problem with the droplet generation. Next, the 2D amplitude plot was opened and each set of wells was analyzed to set a threshold for the fluorescence. The threshold of fluorescence is set by looking at the droplet separation of positive and negative droplets in the controls.

The 2D amplitude shows the droplets for the two sets of multiplex PCRs. The fluorophore for each multiplex PCR is different therefore, the fluorescence data is observed in two channels. Each quadrant shows a different DNA species. On the left topside, Q1 shows the channel 1 positive droplets only (3’ deleted and hypermutated), the right topside, Q2 shows droplets positive for both channel 1 and channel 2 (intact), the left bottom Q3 shows negative droplets and the right bottom Q4 shows the droplets positive for channel 2 only (5’ deleted) as seen in Figure 4.

Due to different florescence characteristics, GAG-ENV-Hypermut and RPP30 Shear/RPP30 have different threshold settings. The data were saved in a csv format for the calculations to determine the intact, 3’ defective/hypermutated, 5’ defective and total HIV-1 DNA copies/ million PBMCS.

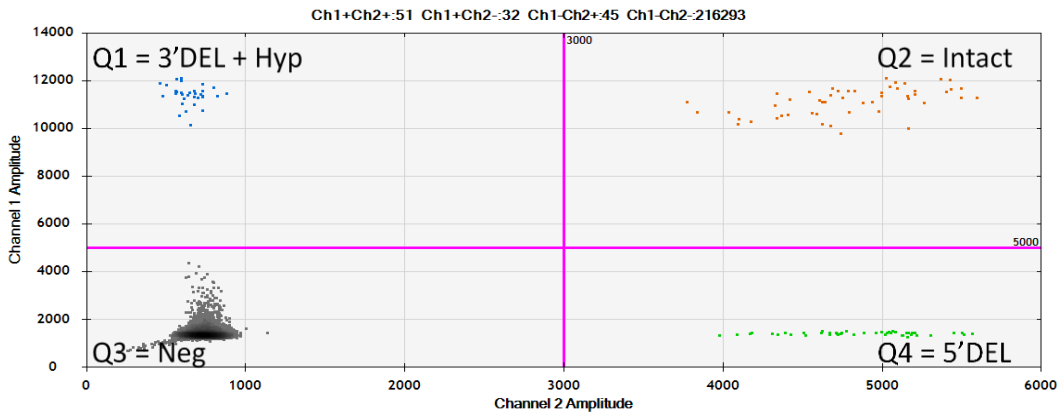


Figure 4. 2D amplitude in Bio-Rad Quantasoft showing droplets.

The 2D amplitude of the IPDA data shows the droplets positive for each PCR reaction in different quadrants. Q1 shows droplets that are positive for GAG amplicon only; Q2 shows droplets that are positive for both GAG and ENV amplicons; Q3 shows droplets that are negative for both GAG and ENV amplicons; Q4 shows droplets that are positive for the ENV amplicon only.

The calculations were done using the formula:

$$\text{Concentration} = -\ln\left(\frac{N_{neg}}{N}\right) \div V_{droplet}$$

where $V_{droplet}$ = volume of a droplet which is 0.85nL, N_{neg} = the number of negative droplets and N = total number of droplets. This formula is given by Bio-Rad Industries to calculate the concentration of HIV-1 DNA species. The formulae based on this were used for the calculations as follows:

Concentration of gag/µL (G):

$$-\ln\left(1-\left(\left(\frac{(N_{Q2}+N_{Q1})}{N}\right)\div\left(\frac{V_{droplet}}{1000}\right)\right)\right)$$

Concentration of env/µL (E):

$$-\ln\left(1-\left(\left(\frac{(N_{Q2}+N_{Q4})}{N}\right)\div\left(\frac{V_{droplet}}{1000}\right)\right)\right)$$

Intact copies/µL (I) =

$$-\ln\left(1-\left(\left(\frac{(N_{Q2})}{N}\right)\div\left(\frac{V_{droplet}}{1000}\right)\right)\right)$$

5' defective HIV- 1 DNA copies/µL (D5) =

$$-\ln\left(1-\left(\left(\frac{(N_{Q4})}{(N_{Q1}+N_{Q3}+N_{Q4})}\right)\div\left(\frac{V_{droplet}}{1000}\right)\right)\right)$$

3' defective/hypermutated HIV- 1 DNA copies/µL (D3) =

$$-\ln\left(1-\left(\left(\frac{(N_{Q1})}{(N_{Q1}+N_{Q3}+N_{Q4})}\right)\div\left(\frac{V_{droplet}}{1000}\right)\right)\right)$$

where: N_{Q2} = number of double positive droplets; N_{Q1} = number of single positive droplets for 3' defective; N_{Q3} = number of negative droplets; N_{Q4} = number of single positive droplets for 5' defective; and $V_{droplet}$ = volume of each droplet i.e., 0.85nL

For copies/well, the copies/µL was multiplied by 20 since that was the volume of the samples used for droplet generation.

In a cell, it is expected that the viral genome would unsheared, but when cells are processed through DNA isolation and ddPCR, shearing of the genomic DNA occurs which can lead to underestimation of the number of intact proviruses. The shearing occurs due to mechanical stresses such as freeze thaw cycles and the passage of DNA through silicon columns in the DNA isolation protocol. This can be accounted for in the final calculations by using the housekeeping gene RPP30. The primers and probes for RPP30 and RPP30 shear bind at a distance of 7kb on the RPP30 gene therefore if this gene shows

shearing it is expected that the HIV-1 genome between psi (upstream of gag) and env is also sheared and to the same extent. The DNA shearing was calculated as follows:

Concentration of $RPP30_1$/µL (R1) =

$$-\ln\left(1-\left(\left(\frac{(N_{Q2}+N_{Q1})}{(N_{Q1}+N_{Q2}+N_{Q3}+N_{Q4})}\right)\div\left(\frac{V_{droplet}}{1000}\right)\right)\right)$$

Concentration of $RPP30_2$/µL (R2) =

$$-\ln\left(1-\left(\left(\frac{(N_{Q2}+N_{Q4})}{(N_{Q1}+N_{Q2}+N_{Q3}+N_{Q4})}\right)\div\left(\frac{V_{droplet}}{1000}\right)\right)\right)$$

%unsheared (S) = $N_{Q2} \div \left(\left(\frac{N_{Q1}+N_{Q4}}{2}\right)+N_{Q2}\right)$

where: N_{Q2} = number of double positive droplets; N_{Q1} = number of single positive droplets for RPP30 shear; N_{Q3} = number of negative droplets; N_{Q4} = number of single positive droplets for RPP30 and $V_{droplet}$ = volume of each droplet, i.e., 0.85nL

Since RPP30 is diluted by a factor of 100, to calculate undiluted cells the following formula was used:

Undiluted cells/µL (U) = $S \div \left(\left(\frac{R_1+R_2}{2}\right)\times 100\right)$

where S = %unsheared; R_1 = concentration of $RPP30_1$; R_2 = concentration of $RPP30_2$.

For undiluted cells/well, the cells/μL was multiplied by 20 since that was the volume of the samples used for droplet generation.

Total cells analyzed (T) = $U \div (number\ of\ wells\ that\ passed\ analysis\ per\ sample)$

where U = Undiluted cells/μL.

To adjust for shearing:

Copies/shear adjusted well for each DNA species = $(copies/well) \div (S)$

where S= %unsheared.

Copies/shear adjusted cells for each DNA species =

$$(copies/shear\ adjusted\ well) \div (U)$$

where U = Undiluted cells/μL.

Copies/10^6 cells shear adjusted for each DNA species =

$$(copies/shear\ adjusted\ cells \times 1000000)$$

Total DNA = $(intact\ cpm + 3'del\ cpm + 5'del\ cpm)$

where cpm = copies/million shear adjusted cells

5. Whole Genome Amplification and Primer Sequencing

Whole genome amplification (WGA) was performed using the Qiagen REPLI-g single cell kit (#150343; Qiagen, Valencia, California) on samples that did not show any result in the IPDA. It was suspected that these samples had primer-probe mismatches[171]. Since DNA was limited, this approach was used to ensure favorable outcomes.

5.1 Whole Genome Amplification

The buffer DLB was reconstituted per the manufacturer's instructions. Buffers D1 and N1 were prepared for 8 samples using the following table:

Component	Buffer D1 x9	Buffer N1 x9
Reconstituted Buffer DLB	63µL	-
Stop Solution	-	81µL
Nuclease Free Water	225µL	459µL

Table 3 (A): Buffer preparation for Whole Genome Amplification

The master mix was prepared using the following table, but polymerase was not added at this stage.

Component	Volume (µL)
Nuclease Free Water sc	61
REPLI-g sc reaction Buffer	261
REPLI-g sc DNA polymerase	18
Total Volume	340

Table 3 (B): Preparation of Master Mix for Whole Genome Amplification

2.5µL of template DNA was added to each of the wells in a fresh 0.2mL 24 well PCR plate followed by the addition of 2.5µL of Buffer D1. The plate was vortexed and spun at 1000 rpm for 15 seconds followed by incubation at room temperature for 3 minutes. The plate was kept on ice, and 5µL of Buffer N1 was added to the wells. The content of the wells was mixed by pipetting up and down. The polymerase was added to the master mix as per the table. The Master mix tube was then vortexed and spun briefly. 40µL of this Master Mix was then added to each of the wells (containing 10µL of denatured DNA) while the plate was still on ice. The plate was vortexed, spun at 1000 rpm for 15 seconds and placed in the thermocycler.

The following program was run:

Temperature	Duration
30°C	8h
65°C	3min
4°C	∞

5.2 PCR for Psi and Env regions

To start sequencing, the WGA products had to undergo two PCR reactions that would ensure that the Psi (upstream of GAG) and Env regions were amplified.

Assay	Primer name	PCR reaction	Primer type	Primer sequence
u5-gag	u5gagSGS_fo	outer	forward	5' GTARCTAGAGATCCCTCAGAC 3'
	u5gagSGS_ro		reverse	3' TGACATGCTGTCATCATYTCYTC 5'
	u5gagSGS_fn	nested	forward	5' AAATCTCTAGCAGTGGCGCC 3'
	u5gagSGS_rn		reverse	3' CATCATTTCTTCTARTGTAGCTSCT 5'
Env	envSGS_fo	outer	forward	5' GCCAGTAGTRTCAACYGAA 3'
	envSGS_ro		reverse	3' GCARATGAGTTTCTYAGAGCA 5'
	envSGS_fn	nested	forward	5' CTGCTAAATGGCAGTCTAGC 3'
	envSGS_rn		reverse	3' TTGCCTGGAGCTGYTTRATGC 5'

Table 4: Primers for Psi and Env PCR[171]

The next day, the Master Mix for the outer PCR was prepared using the table below:

Components	µL/RXN	Total µL
10x PCR buffer (Invitrogen; #P/N 52045)	2.5	22.5
50mM MgCl2 (Invitrogen; # P/N52044)	1	9
10mM PCR Nucleotide mix (Promega; #U151B)	0.5	4.5
50µM Forward primer Outer	0.5	4.5
50µM Reverse primer Outer	0.5	4.5
5U/µL Platinum Taq (Invitrogen; #11304-011)	0.2	1.8
Nuclease Free Water	19.8	178.2
Total Vol	**25**	**225**
Template Vol	**1**	**9**

Table 5 (A): Preparation of Master Mix for Psi-Env Outer PCR

2µL of the WGA product was taken and added to a fresh 0.2mL 24 well PCR plate. 8µL of the Master Mix for Outer PCR was added. The plate was vortexed, spun at 1000 rpm for 15 seconds and placed in the thermocycler[171]. The following program was run:

Temperature	Duration	
94°C	2min	
94°C	30sec	x40 Cycles
50°C	30sec	
72°C	90sec	
72°C	3min	
4°C	∞	

After this the outer PCR product was taken and 10µL of Tris-HCl 10mM was added to it to dilute the outer PCR product for a nested PCR. The Master Mix for the nested PCR was prepared using the table below:

Components	µL/RXN	Total µL
10x PCR buffer (Invitrogen; #P/N 52045)	2.5	22.5
50mM MgCl2 (Invitrogen; # P/N52044)	1	9
10mM PCR Nucleotide mix (Promega; #U151B)	0.5	4.5
50µM Forward primer Nested	0.5	4.5
50µM Reverse primer Nested	0.5	4.5
5U/µL Platinum Taq (Invitrogen; #11304-011)	0.2	1.8
Nuclease Free Water	19.8	178.2
Total Vol	**25**	**225**
Template Vol	**1**	**9**

Table 5 (B): Preparation of Master Mix for Psi-Env Nested PCR

2µL of the diluted outer PCR product was taken and added to a fresh 0.2mL 24 well PCR plate. 18µL of the Master Mix for Nested PCR was added. The plate was vortexed, spun at 1000 rpm for 15 seconds and the following program was run:

Temperature	Duration	
94°C	2min	
94°C	30sec	x40 Cycles
55°C	30sec	
72°C	90sec	
72°C	3min	
4°C	∞	

The PCR product was run on a gel to determine the correct band size. If the band was of the appropriate size, the product was tip purified and sent for Sanger Sequencing[171]; else, the sample was kept aside for the near full-length genome PCR.

5.3 Four fragment PCR and Alternative env PCR

The primer sets for each fragment are listed in Table 6.

Fragment	Primer set (Forward/Reverse)
A	275F/3 In Out
B	263 F/ 3 AccOut
C	5 In out/BLInnerR
D	5 AccOut/280R

Table 6: Primers for the near-full length Four fragment PCR

For the samples that did not show bands in the Gels for the nested u5gag and env PCR, a near full-length four fragment PCR was run. The outer PCR master mix was prepared using the table below:

REAGENT	Volume per RXN (µL)	# of RXNS	Master Mix (µL)
5X LongAmp Buffer	5	6	30
dNTP (10mM)	0.75	6	4.5
BLouterF Primer (10 µM)	1	6	6
BLouterR Primer (10 µM)	1	6	6
H_2O	14.58	6	87.48
LongAmp Hot Start Taq DNA Polymerase	1	6	6
Total	23.33		139.98
Template DNA	1.67		

Table 7: Preparation of Master Mix for Four fragment outer PCR

1.67µL of template DNA from the samples (except sample 3 at W208, there was not enough DNA, so the WGA product was taken), was added to a fresh

0.2mL 24 well PCR plate and followed by 23.33µL master mix. The plate was vortexed, spun at 1000 rpm for 15 seconds and the following program was run:

Temperature	**Duration**	
94°C	2min	
94°C	30sec	X3 Cycles
64C	30sec	
68°C	10 min	
94°C	30 sec	X4 Cycles
61°C	30 sec	
68°C	10 min	
94°C	30sec	X3 Cycles
58°C	30sec	
68°C	10 min	
94°C	30 sec	X36 Cycles
55°C	30 sec	
68°C	10 min	
68°C	10 min	
4°C	∞	

The next day, the outer PCR product was diluted 1:3 by adding 50 µL of Nuclease free water to the wells. The plate was vortexed and spun at 1000 rpm for 15 seconds.

The master mix for each of the fragments was prepared using the tables below:

Reagent	Volume per RXN (µL)	# of RXNS	Master Mix (µL)
5X LongAmp Buffer	10	5	50
dNTP (10 uM)	1.5	5	7.5
Forward primer (10 uM)	2	5	10
Reverse primer (10 uM)	2	5	10
H_2O	29.16	5	145.8
LongAmp Hot Start Taq	2	5	10
Template DNA	**3.34**		
TOTAL	**46.66**		233.3
For Fragment C & D + alternative ENV primers, 6 rxns were prepared			

Table 8: Preparation of Master Mix for Four fragment inner PCR

3.34µL of template DNA from the samples, was added to a fresh 0.2mL 24 well PCR plate and followed by 46.66 µL master mix. The plate was vortexed, spun at 1000 rpm for 15 seconds and the following program was run:

Temperature	Duration	
94°C	2min	
94°C	30sec	X3 Cycles
64C	30sec	
68°C	5/6:30 min *	
94°C	30 sec	X4 Cycles
61°C	30 sec	
68°C	5/6:30 min	
94°C	30sec	X3 Cycles
58°C	30sec	
68°C	5/6:30 min	
94°C	30 sec	X36 Cycles
55°C	30 sec	
68°C	5/6:30 min	
68°C	5/6:30 min	
4°C	∞	

*the time varied based on primers for each fragment. The 68°C step for outer PCR for Fragment A & D was run for 5 min and for Fragment B & C it was run for 6 min 30 seconds.

The PCR product was run on a gel to determine correct band size. If it matched, the product was tip purified and sent for Sanger Sequencing[171].

6. HIV-1 Parameters and Characteristics for the PHACS-AMP Cohort

In the current study, remnant genomic DNA samples from a prior study on the Adolescent Master Protocol (AMP), the Pediatric HIV AIDS Cohort Study (PHACS) were utilized. In the prior study, a singleplex HIV-1 Pol ddPCR assay

was used for samples from the PHACS-AMP cohort to study HIV-1 reservoir dynamics in perinatal infections[60]. PHACS is a longitudinal study based in the United States to understand the long-term outcomes of perinatal HIV-1 infections[210]. In PHACS, over 451 children/adolescents living with perinatal HIV-1 infection were enrolled of which 61 were eligible for the previous study based on virologic control and number of available samples[60]. The study reported here was further limited based on having sufficient available genomic DNA, and detectable levels of HIV-1 DNA upon testing using the POL-LTR ddPCR assay[60,145]. Since the present study was going to be run on remnant samples from previous studies done on the same cohort in the lab[60], a few criteria were used to determine which samples were to be tested. The criteria were as follows:

- The remnant sample needed to have a minimum of 2000ng of DNA to allow for sufficient input genomic DNA[60].
- Samples needed to have 2LTR circles ≤ 10 copies/million PBMCs, to be in the more steady state levels of reservoirs on effective ART.
- The participant should have been on effective ART with undetectable plasma viral loads in clinical assays, i.e., achieved virologic suppression (as defined below).

Virologic suppression was defined as plasma viral load (pVL) ≤400 copies/mL for the duration of the study while permitting single pVL≥400 copies/mL in between measurements of pVL≤400 copies/mL[59,60]. Of the 61 participants in the previous study, 25 matched these criteria. Their HIV parameters and characteristics are presented in table 9.

		Age at virologic suppression	
	Total (N=25)	*<1 year old (N=8)*	*1 to 5 years old (N=17)*
Sex			
M	10 (40%)	3 (38%)	7 (41%)
F	15 (60%)	5 (63%)	10 (59%)
Year of birth			
Mean (SD)	1997 (3)	1998 (1)	1997 (3)
Median	1997	1998	1996
Q1, Q3	1995, 1999	1997, 2000	1994, 1999
HIV prophylaxis[$]			
Yes	5 (20%)	2 (25%)	3 (18%)
No	20 (80%)	6 (75%)	14 (82%)
Number of pre-cART[#] ART regimens			
Mean (SD)	0.76 (1.09)	0.13 (0.35)	1.06 (1.20)
Median	0	0	1
Q1, Q3	0, 1	0, 0	0, 2
Duration of pre-cART[#] ART regimens (years)			
Mean (SD)	0.77 (1.25)	0.00 (0.01)	1.13 (1.38)
Median	0.00	0.00	0.42
Q1, Q3	0.00, 1.19	0.00, 0.00	0.00, 1.66
Type of cART[+]			
PI alone	18 (72%)	5 (63%)	13 (76%)
PI + NNRTI	4 (16%)	1 (13%)	3 (18%)
NNRTI alone	3 (12%)	2 (25%)	1 (6%)
Age at cART[+] initiation (years)			
Mean (SD)	1.73 (1.70)	0.24 (0.11)	2.44 (1.64)
Median	1.13	0.24	2.23
Q1, Q3	0.25, 3.33	0.15, 0.28	1.13, 3.60
Age at cART[+] initiation (months)			
Mean (SD)	20.80 (20.42)	2.87 (1.30)	29.25 (19.70)

		Age at virologic suppression	
	Total (N=25)	*<1 year old (N=8)*	*1 to 5 years old (N=17)*
Median	13.60	2.89	26.81
Q1, Q3	2.96, 40.01	1.77, 3.38	13.60, 43.20
HIV RNA (log_{10} c/mL) at cART^{+} initiation			
Mean (SD)	4.82 (0.87)	4.76 (1.36)	4.85 (0.50)
Median	4.90	5.23	4.83
Q1, Q3	4.30, 5.43	3.54, 5.82	4.38, 5.35
Missing	2	0	2
CD4% at cART^{+} initiation			
Mean (SD)	30 (11)	37 (9)	26 (10)
Median	30	33	28
Q1, Q3	25, 39	31, 42	23, 33
Missing	2	0	2
Age at confirmed virologic suppression* (years)			
Mean (SD)	2.5 (1.7)	0.66 (0.2)	3.4 (1.4)
Median	2.2	0.66	3.7
Q1, Q3	0.90, 3.8	0.46, 0.86	2.2, 4.7
Age at confirmed virologic suppression* (months)			
Mean (SD)	30.3 (21.2)	7.9 (2.5)	41.6 (16.9)
Median	26.9	7.9	44.3
Q1, Q3	10.7, 45.9	5.4, 10.2	26.9, 56.9
Years from cART initiation to confirmed virologic suppression*			
Mean (SD)	0.84 (1.0)	0.42 (0.1)	1.0 (1.2)
Median	0.44	0.36	0.46
Q1, Q3	0.32, 0.59	0.30, 0.57	0.33, 1.3
Months from cART initiation to confirmed virologic suppression*			
Mean (SD)	10.0 (12.7)	5.0 (2.1)	12.3 (14.9)
Median	5.2	4.3	5.5
Q1, Q3	3.8, 7.1	3.5, 6.9	3.9, 15.8

		Age at virologic suppression	
	Total (N=25)	*<1 year old (N=8)*	*1 to 5 years old (N=17)*
Age (years) at first specimen			
Mean (SD)	5.0 (2.5)	3.3 (2.5)	5.8 (2.2)
Median	5.1	3.8	6.3
Q1, Q3	3.2, 7.3	0.86, 4.5	4.4, 7.4
Age (years) at last specimen			
Mean (SD)	8.0 (2.3)	6.8 (1.8)	8.5 (2.4)
Median	7.8	6.7	8.7
Q1, Q3	6.3, 9.7	5.4, 8.4	7.3, 10.6

Table 9: HIV parameters and characteristics, by age at virologic suppression

#pre-cART involves the administration of ARV prophylaxis

+cART = combination antiretroviral therapy

*confirmed virologic suppression is defined as plasma viral load (pVL) ≤400 copies/mL[59,60]

7. Statistical Analysis

Optimizing the assay:

The intact HIV-1 DNA copies/million PBMCs for each amount of DNA input were run against the JLAT standards on an XY plot followed by analysis using simple linear regression in GraphPad Prism v8.4.3. The %unsheared data sets were analyzed using non-parametric independent t-tests (Mann-Whitney) and with significance defined as $\alpha = 0.05$.

For the purposes of this study, the limit of detection (LOD) was calculated using the number of cells analyzed and therefore varied between samples. For each DNA input, the LOD of the dataset was determined by the lowest copy number of HIV-1 DNA concentration that was detectable in >95% of the runs.

Applying the assay to participant samples:

The intact, 5'deleted, 3'deleted/hypermutated and total HIV-1 DNA copies/million PBMCs were analyzed using box plots, and non-parametric unpaired Mann Whitney t-tests were used to compare runs between the four HIV-1 DNA species with significance defined as $\alpha = 0.05$. The DNA species were analyzed as mentioned above for both age at virologic suppression and duration of virologic suppression. As the laboratory members were blinded to the distribution of samples by age at virologic suppression, the PHACS statisticians performed the analysis for age at virologic suppression using the descriptive statistics function in SAS v9.4. The p-values for this stratification were not calculated owing to the small size of the cohort.

The rest of the analysis was done by the author of the thesis using GraphPad Prism v8.4.3. The number of cells analyzed per participant/time point and the % unsheared per participant/time point were plotted and non-parametric unpaired Mann Whitney t-tests were run with significance defined as $\alpha = 0.05$. The % intact for each time point was calculated by dividing the intact copies/million PBMCs by the total copies/million PBMCs, and non-parametric unpaired Mann Whitney t-tests were run between the three sub-groups of duration of virologic suppression with significance defined $\alpha = 0.05$. The p-values might be affected by the size of the cohort, number of time points per participant and limit of detection for each sample based on the input number of cells analyzed. The total DNA copies/million PBMCs obtained from the POL-LTR and IPDA assays were compared, and a non-parametric Spearman correlation was performed. The same was repeated for each species detected

by IPDA vs the total DNA from the POL-LTR assays. The t-tests were unpaired as not every sample had paired values for all time points. The medians and the interquartile range for all the data were calculated using the descriptive statistics function in GraphPad Prism v8.4.3 and SAS v9.4.

III. Results

1. Optimizing the Intact Proviral DNA Assay (IPDA)

1.1. Establishing controls for the IPDA.

Cell lines were tested as controls. JLATs were used as they contain one copy of integrated HIV-1 DNA per cell, which is thought to be the case for HIV-1 infection[226]. The JLATs were tested using 1000 ng input of genomic DNA per well in eight replicates to maximize detection at low copy numbers of HIV-1 DNA. As shown in Figure 5A, linearity was observed at higher concentrations, and down to 25 HIV-1 DNA copies/million PBMCs between the expected and observed intact HIV-1 DNA from the JLAT cells for all three DNA input amounts. This showed that the dilution series worked well with the assay at concentrations >25 copies/million PBMCs. At 10 HIV-1 DNA copies/million PBMCs, we started losing linearity, and precision and the variation was increasing perhaps owing to Poisson statistics. This implied that to gain more precision at and below 10 HIV-1 DNA copies/million PBMCs, more experiments needed to be run with higher number of replicates.

Based on the results, it was concluded that JLAT cells performed well in terms of linearity and quantification compared to the expected HIV-1 DNA copies/million PBMCs. When studying a wide range of proviral loads, JLAT cells can inform about assay characteristics over this range and could be used to

further define the assay characteristics in lower range of HIV-1 proviral DNA copies detected in paediatric samples.

1.2. Sensitivity of IPDA

The next step was to determine the minimal DNA input at which the assay can successfully detect intact HIV-1 DNA copies using different input genomic DNA, since it may not always be possible to have sufficient DNA in the samples to input 1000ng in eight replicates. Therefore, the chosen input genomic DNA amounts were added at 500ng and 200ng. As seen in Figure 5B and C, there was good linearity between the expected and observed intact HIV-1 DNA copies/million PBMCs at these lower inputs of genomic DNA down to 25 HIV-1 DNA copies/million PBMCs. For 1000ng, the assay could not reliably detect intact copies below a concentration of 5 intact HIV-1 DNA copies/million PBMCs, suggesting a limit of detection of 5 intact HIV-1 DNA copies/million PBMCs. With 200ng and 500ng DNA input, the detection was not reliable below 5 intact copies/million PBMCs.

For 500ng, 10 intact HIV-1 DNA copies/million PBMCs were detected 100% of the time, and 5 and 2.5 intact HIV-1 DNA copies/million PBMCs were detected 75% of the time. For 200ng, 10 intact HIV-1 DNA copies/million PBMCs were detected 100% of the time, and 5 and 2.5 intact HIV-1 DNA copies/million PBMCs were detected 25% of the time.

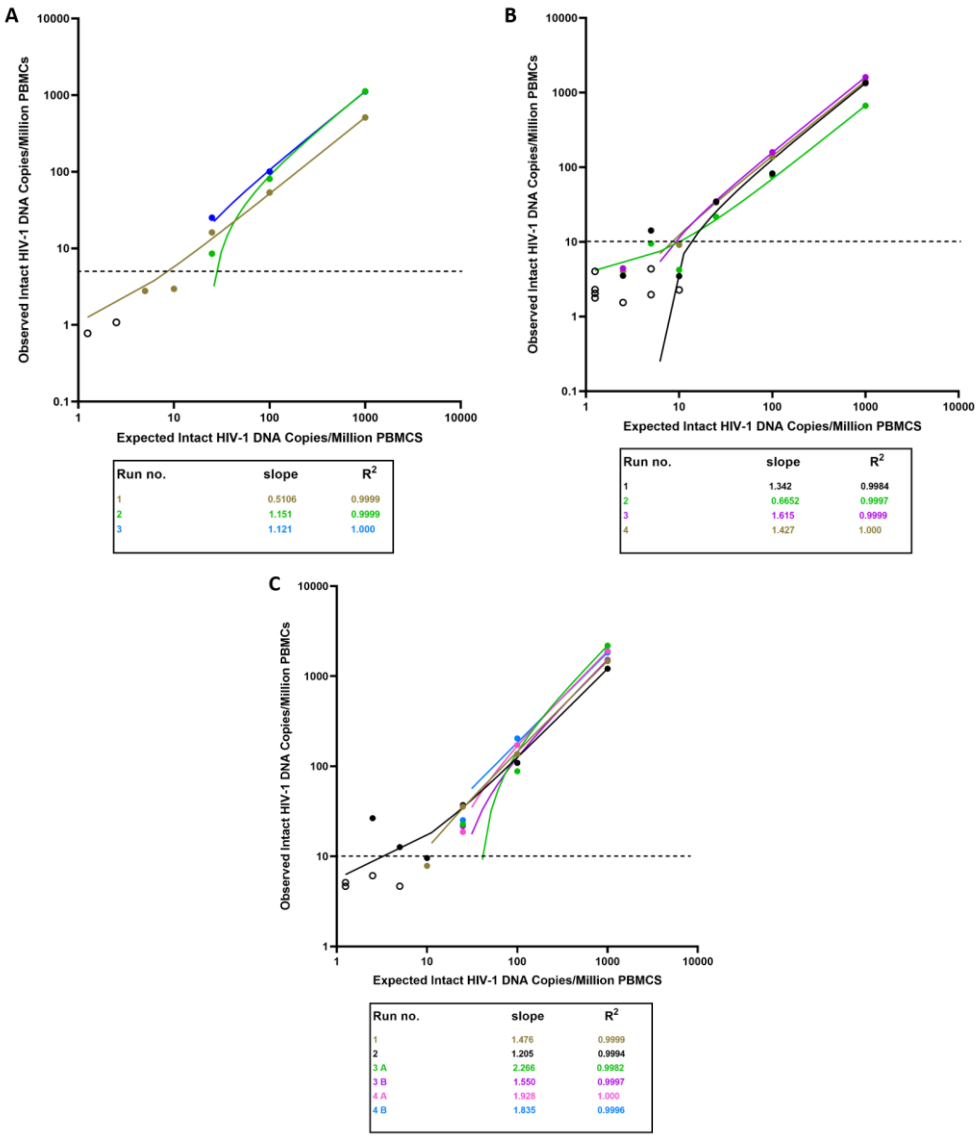


Figure 5. Intact HIV-1 DNA copies/million with different input DNA amounts.

A: Three runs were performed with 1000ng DNA from JLAT+ PBMC standard (two different blood bags). **B:** Four independent runs were performed with DNA isolated from JLAT+ PBMC standards, normalized to a concentration of 500ng (two different blood bags). **C**. Four independent runs were performed with genomic DNA isolated from JLAT+ PBMC standards, normalized to a concentration of 200 ng (one blood bag).

Each run is represented in a different color. Linear regression for all the data points in at the three DNA input concentrations was performed. The linear regression can be observed with the solid line. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. These samples were set to their respective limit of detection values, which vary with the number of cells analyzed. The dotted line indicates the LOD of the assay for each DNA input. Blood bags were obtained from the New York Blood Center and processed as described in methods. Runs 2,3 of 1000 ng and 3,4 of 200 ng had only 3 data points corresponding to I1,I2,I3. Some data points are superimposed in figure 5B & C; however, the number of runs is as stated earlier.

The IPDA was developed for adult samples, where the median proviral load is between 10-100 HIV-1 DNA copies/million PBMCs[69]. Pediatric patients on early long-term suppressive ART tend to have proviral loads <20 HIV-1 DNA copies/million PBMCs[59,60,196,199]. The poor reliability and high variation of the assay below 10 HIV-1 DNA copies/million PBMCs makes it better suited for late suppression than early suppression. However, testing with more replicates may help improve the sensitivity of the assay.

Following testing with different amounts of input DNA, the number of cells analyzed, and the percentage of shearing were the next characteristics that had to be determined. The number of cells analyzed is affected by the number of replicates (wells) and the total amount of DNA input per well. If the amount of DNA input is higher or the number of replicates is higher or both, more cells will be analyzed. It was observed that for 8000ng input DNA for a sample (1000ng in 8 replicates), a median of 697,827 cells were analyzed, while for 4000ng (500ng in 8 replicates) input DNA, a median of 462,810 cells were analyzed, and for 1600ng (200ng in 8 replicates) input DNA, a median of 135,418 cells were analyzed (Data not shown).

Shearing is most heavily affected by the DNA isolation step of the IPDA owing to mechanical stresses, freeze thaw cycles and the silicon column used for isolation. For the standards, the Qiagen Qi-Amp Midi kit was used to isolate DNA and then the IPDA was run. It was noted that the percent unsheared was highly variable and ranged from 35-75%, with the majority of samples lying in between 50-70% (Figure 6). The median for the 85 data points analyzed was 62% unsheared, with the interquartile range of 57.3%- 64.4%.

Although the %unsheared was similar to the previously published range as reported[69], it was important to minimize shearing as much as possible since pediatric participant samples can have exceedingly low concentrations of HIV-1 DNA and therefore low % unsheared could cause the assay to miss the intact provirus in the participants.

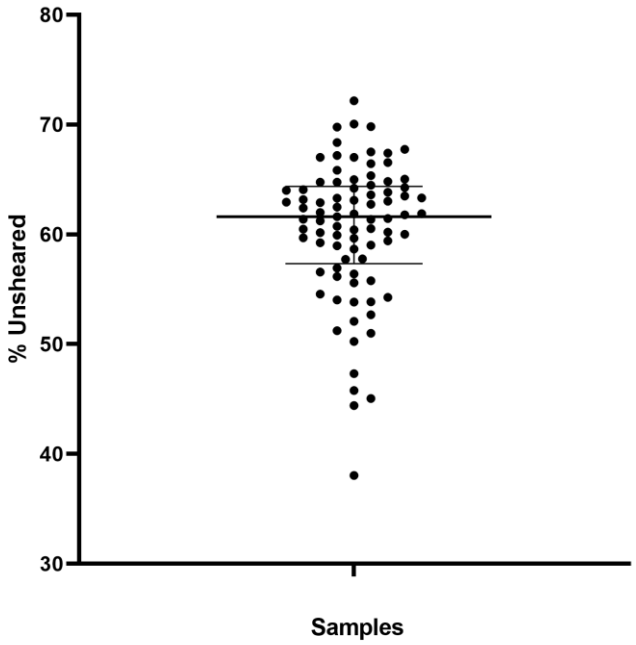


Figure 6. Range of % Unsheared observed over a series of runs.

The range of % unsheared HIV-1 DNA observed over different experiments, was mostly concentrated between 50 – 70% using the Qiagen Qi-AMP Blood DNA Midi kit. The median was 61.6% for n=85 samples. The interquartile range (IQR) was 57.3% - 64.4%.

1.3. Impact of method of DNA isolation on Shearing

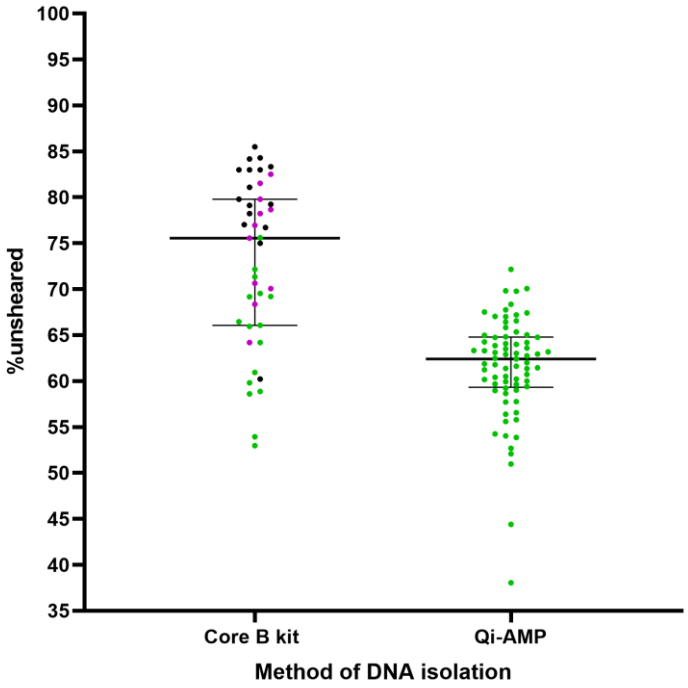
Paediatric samples have low proviral loads[59,60,196,199]. Therefore, we strived to achieve as high a %unsheared as possible and since reducing shearing is high priority, we tested out a different method of isolation using the Qiagen Gentra

Puregene Blood kit (Cat # 158467). As seen in Figure 7, the % unsheared on average was higher for Qiagen Gentra Puregene Blood kit compared to the Qiagen Qi-Amp Blood DNA Midi kit, and the difference was statistically significant, $p<0.0001$. Over three performers, lower shearing was observed using the Gentra Puregene protocol vs the Qi-amp that was performed by one person. Analysis of the other samples isolated using the Qiagen Qi-Amp kit by other performers in our laboratory have shown higher shearing percentages (data not shown).

The Qiagen Gentra Puregene Blood kit performed better since it does not include columns, which can enhance the shearing of DNA. Hence, Qiagen Gentra Puregene Blood kit is better suited for isolating genomic DNA for assays that require higher concentration of unsheared DNA, like the IPDA. However, this was not possible for the PHACS samples used in this study as the DNA was previously isolated for another study[60] using the Qiagen Qi-AMP Blood DNA Midi Kit (Cat #51183).

2. Applying the IPDA to PHACS samples

Next, the IPDA was applied to study the reservoir dynamics in participant samples from a previous study in which we observed differences in proviral DNA decay as a function of age at virologic suppression[60]. However, the prior study was limited by the single-plex HIV-1 DNA assay used to measure total proviral reservoir size because that assay did not distinguish intact from defective proviruses.



N	43	73
Median	75.6	62.4
(IQR)	66.1 - 79.8	59.3 - 64.8

Figure 7. % Unsheared observed from DNA isolated using two different methods and kits.

The different colors represent DNA isolated from different technicians. The green symbols indicate DNA isolated by Performer 1, purple symbols indicate DNA isolated by Performer 2 and the black symbols indicate DNA isolated by Performer 3.

2.1 Participant characteristics

As mentioned earlier, in this study, DNA samples from the previous study on the PHACS-AMP cohort were used. 25 participants matched the criteria for the current study mentioned earlier. Of the 25 participants, 8 had achieved virologic suppression (VS) before 1 year of age (median = 0.66 years), while 17 had achieved virologic suppression between 1-5 years of age (median = 3.4 years),

as seen in Table 9. The median age at confirmed VS was 7.9 months for the group that achieved VS<1 year of age (group 1) and 44.3 months for those that achieved VS between 1-5 years of age (group 2). The ratio of females to males studied was 3:2 for the entire cohort (Table 9).

Out of the 25 participants, 1 had six time points available for study, 4 had five, 5 had four, 3 had three, 7 had two and 5 had one, for a total of 74 time points available for study. However, owing to our selection criteria mentioned in the methods, time points that had 2LTR ≥ 10 (5 time points) and purported primer-probe mismatch issues (7 time points) were excluded from the analysis. One participant had a time point which had to be eliminated due to extremely low droplet formation in the IPDA. Hence of the original 74 time points, only 61 were used for analysis.

2.2 Performance characteristics of the IPDA on PHACS-AMP

Some of the most important performance characteristics that need to be analyzed for the IPDA are the % unsheared and the number of cells analyzed, as these affect the conclusions that can be derived from the assay. The % unsheared for this study set (Figure 8) ranged between 40 – 55% with a median of 46%, which is lower than that observed for the standards (Figure 6) that ranged primarily between 50 - 70% with a median of 61.6%. Since the study used samples whose DNA was isolated 5 years ago, tested multiple times and underwent several rounds of freeze thawing, relatively low % unsheared was expected.

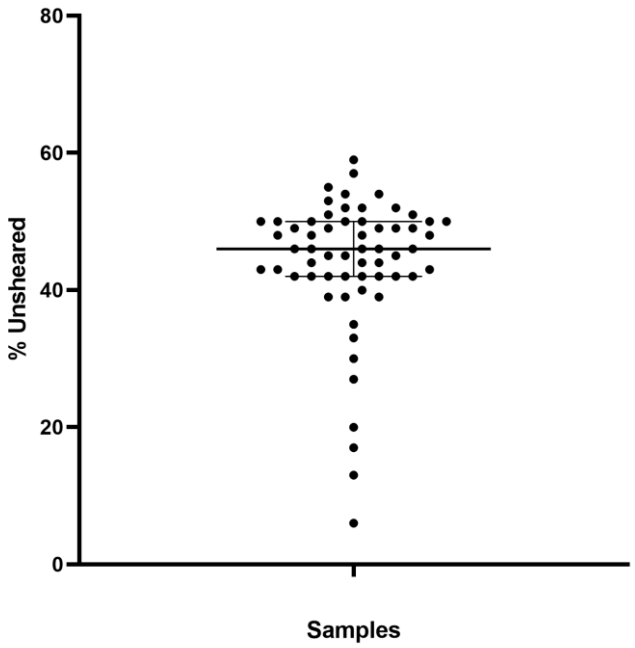


Figure 8. % Unsheared for the cohort.

The range of % unsheared observed over different samples was mostly concentrated between 40 – 55%. The % unsheared contains data from participants at most times points. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues are excluded. N= 61, Median %unsheared = 46, Interquartile range (IQR) = 42 – 50. The DNA for these samples was isolated over 5 years ago and has been tested several times and underwent various mechanical stresses which affect the shearing.

With respect to the number of cells analyzed, this ranged from 100,000 to 360,000 cells, with a median of 232,384 cells and with most of the samples lying between 200,000 - 270,000 (Figure 9A). The cells analyzed/well ranged from 15,000 - 45,000 cells/well with a median of 29,048 cells analyzed/well, and with most lying between 25,000 - 35,000 cells/well (Figure 9 B). Another metric to gain insight into the number of cells analyzed is cells analyzed/well. The number of total cells analyzed is affected by the number of wells excluded from analysis i.e., with more wells analyzed, the total number of cells will increase

and vice versa. Generally, the expected range of cells analyzed for 200ng DNA input in 8 replicates is between 250,000 - 300,000 cells. But to study and fully characterize the reservoir more cells along the lines of 800,000 - 1,000,000 should be analyzed.

2.3 Data Analysis by Stratification

For this study, the data were stratified by two factors: age at virologic suppression (VS) and duration of virologic suppression. Based on these factors, two and three subgroups were defined as follows:

Age at virologic suppression: <1 year (21 timepoints on 8 participants); 1-5 years (40 timepoints on 17 participants)

Duration of virologic suppression: <2 years (22 timepoints), 2-5 years (34 timepoints), >5 years (6 timepoints)

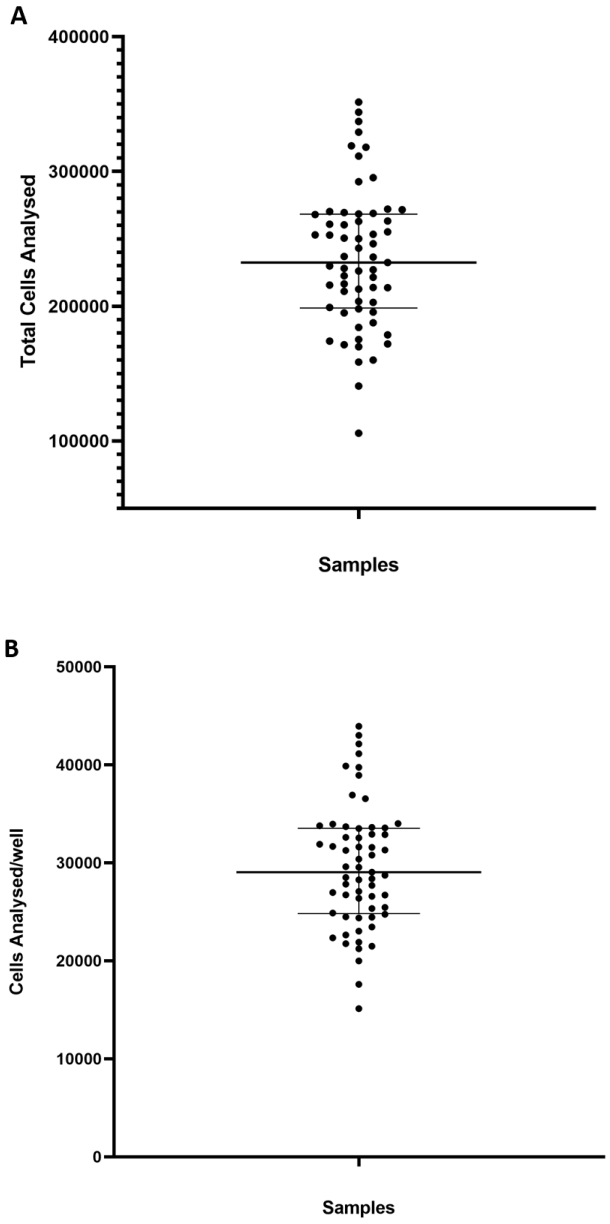


Figure 9. Total cells analyzed and cells analyzed/well for the cohort.

Panel A: Total cells analyzed. The total cells analyzed were mostly concentrated between 200000 – 270000. Total cells analyzed: N= 61, Median = 232384, Interquartile range (IQR) = 198531 – 268241

Panel B: Cells analyzed per well. The cells analyzed/well were mostly concentrated between 25000 - 35000. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The total cells analyzed by the assay are impacted by the replicate wells. Cells analyzed/well: N= 61, Median = 29048, Interquartile range (IQR) = 24816 – 33530.

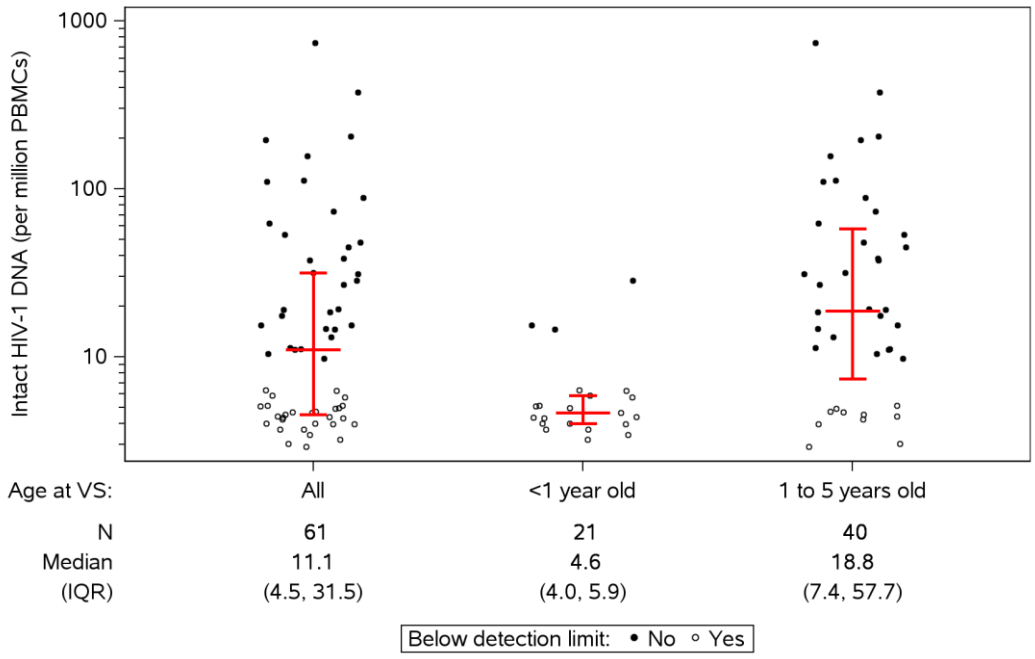


Figure 10. Distribution of PBMC-associated intact HIV-1 DNA.

The plot is based on age at virologic suppression. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range VS: Virologic Suppression.

Overall, intact HIV-1 proviruses were detected in 33/61 samples (54.1%) tested on the 25 participants (8 with VS<1 year of age and 17 with VS between 1-5 years of age) studied with an input concentration of 200 ng in 8 replicates and a median of 232,384 cells analyzed (IQR 198,531 - 268,241). The median concentration of intact HIV-1 proviruses detected was 11.1 copies per million PBMCs (range 4.5 - 57.7 copies/million PBMCs). When stratified by age at virologic suppression (VS), it was observed that group 1 (age at VS<1) had fewer detectable intact HIV-1 DNA copies/million PBMCs (3/21) than group 2 (age at VS between 1-5) which had 30/40 intact proviruses detectable.

When stratified by duration of virologic suppression, it was observed that the intact HIV-1 DNA declined with increasing duration of suppression.

The intact HIV-1 proviral copies (Figure 10), for the participants who achieved VS before 1 year of age, were lower (Median = 4.6 intact HIV-1 DNA copies/million PBMCs) compared to those who achieved VS between 1-5 years of age (Median = 18.8 intact HIV-1 DNA copies/million PBMCs); however, it was important to note that the median of 4.6 HIV-1 DNA copies/million PBMCs for the early suppressed group was not precise since it lay amongst the pool of undetectable proviral loads while the median for the later suppressed group was definitive. For the values that were undetectable, we assigned them the limit of detection of each sample, which were the highest values we could have assigned. Therefore, the median HIV-1 DNA copies/million PBMCs was at most 4.6 for the early suppressed group. Our conservative approach ensured that any significant change we noted with the current analysis would add more confidence that it is a real observable change.

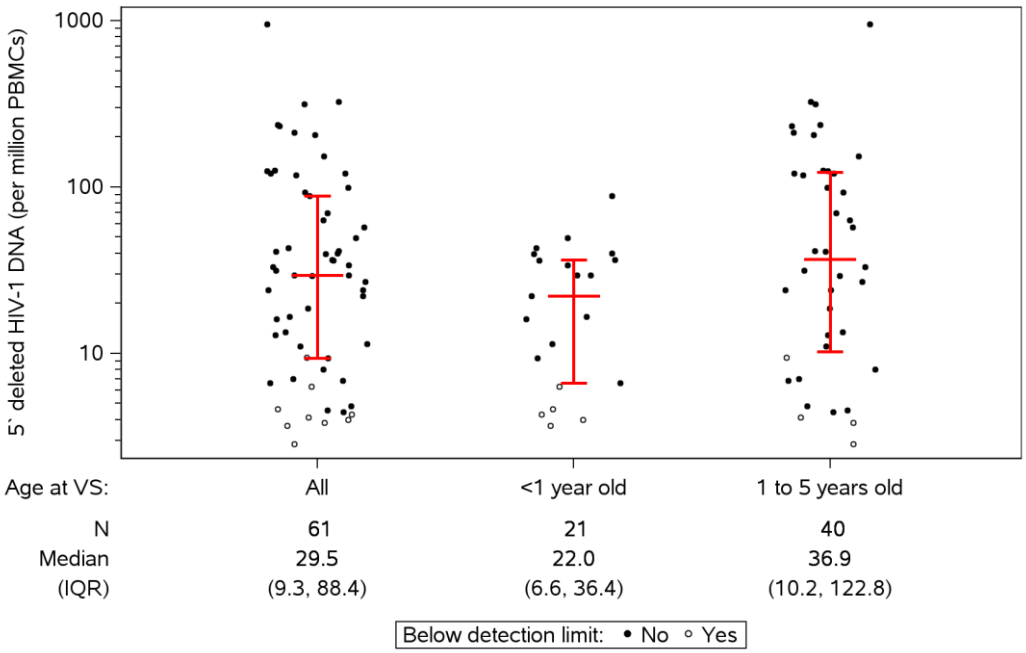


Figure 11. Distribution of PBMC-associated 5'-deleted (defective) HIV-1 DNA.

The plot is based on age at virologic suppression. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range

Similar trends were seen for 5’ deleted HIV-1 proviruses copies (Figure 11) and 3’ deleted/hypermutated HIV-1 proviruses (Figure 12) with medians of 22 and 36.9 5’deleted HIV-1 DNA copies/million PBMCs, and 16.1 and 66.1 3’ deleted/hypermutated HIV-1 DNA copies/million PBMCs for age at VS<1 and age at VS between 1-5 years, respectively.

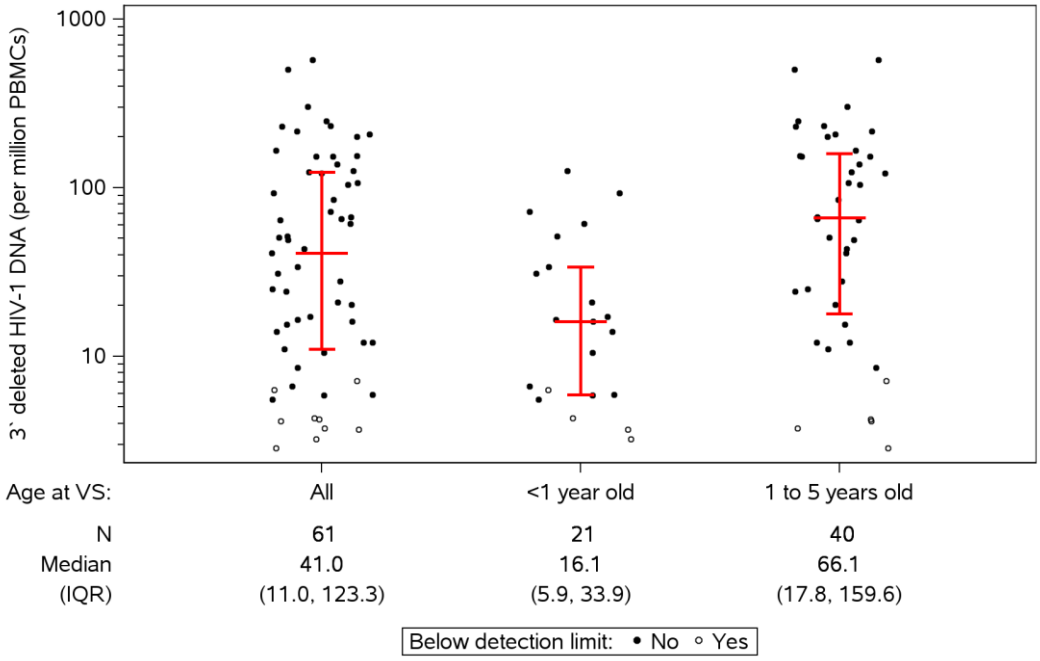


Figure 12. Distribution of PBMC-associated 3'-deleted/hypermutated (defective) HIV-1 DNA.

The plot is based on age at virologic suppression. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range

Since the total DNA is the sum of intact and defective HIV-1 DNA copies/million PBMCs, this trend was also observed for total copies HIV-1 DNA copies/million PBMCs with medians of 32.2 and 178.9 total HIV-1 DNA copies/million PBMCs for the respective age groups (Figure 13).

The data for the intact HIV-1 DNA copies/millions PBMCs for those who achieved VS<1 year of age could have been potentially affected by the lower number of time points (21) in this group compared to the other group (VS 1-5 years of age) which had almost 2X more time points (40). Therefore, by analyzing more time points and increasing the amount of DNA input, it may be

possible to get a better idea of the distribution of intact proviruses in the group that achieved VS<1 year of age.

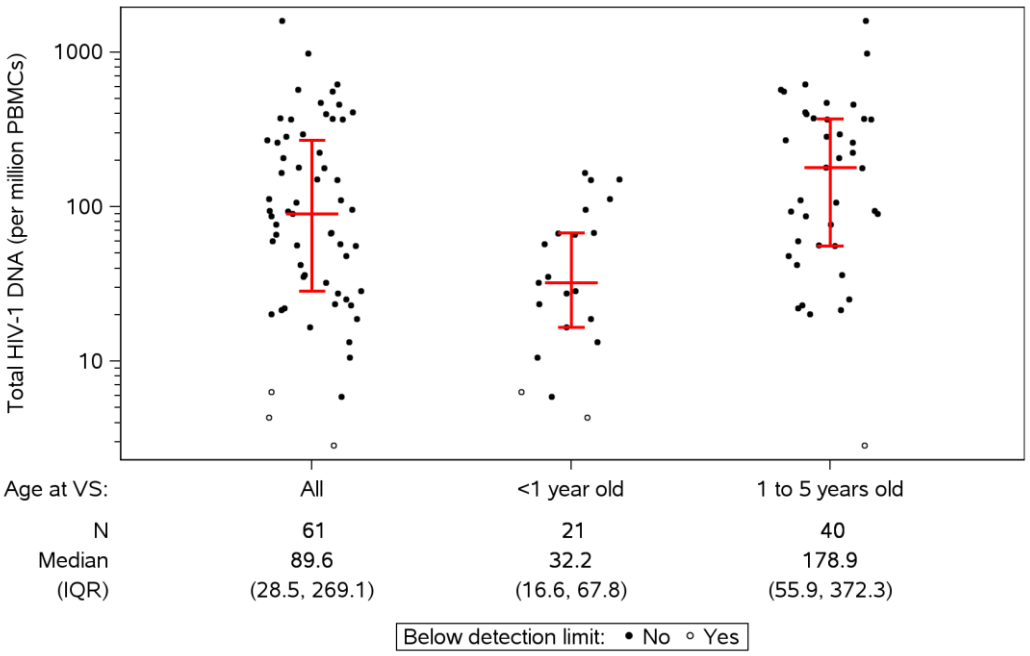


Figure 13. Distribution of PBMC-associated total (intact and defective) HIV-1 DNA.

The plot is based on age at virologic suppression. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range

To understand the trends better and determine if the HIV-1 DNA copies/million PBMCs decayed over time for each group of age at VS, the first and last time points were further analyzed.

The intact HIV-1 DNA copies/million PBMCs seemed to have decreased between the first (Figure 14 A) and last (Figure 14 B) time points for the group that achieved VS<1 year of age with medians of 4.9 and 4.6 intact HIV-1 DNA copies/million PBMCs, respectively. The intact HIV-1 DNA copies/million PBMCs for the group that achieved VS between 1-5 years of age also seemed

to have decreased between the first (Figure 14 A) and last time (Figure 14 B) points with medians of 57.7 and 11.2 intact HIV-1 DNA copies/million PBMCs. However, this trend was not observed for the defective and total (defective + intact) HIV-1 DNA copies/million PBMCs.

In the case of 5' deleted HIV-1 DNA copies/million PBMCs, there was an increase between the first (Figure 15 A) and last (Figure 15 B) time points for both the groups of age at VS with medians of 6.3 and 33.9 (Age at VS<1) and 30.3 and 37.2 (Age at VS 1- 5 years) HIV-1 DNA copies/million PBMCs, respectively.

For the 3' deleted/hypermutated HIV-1 DNA copies/million PBMCs, there was an increase between the first (Figure 16 A) and last (Figure 16 B) time points for both the groups of age at VS with medians of 10.5 and 20.8 (Age at VS<1) and 44.6 and 87.1 (Age at VS 1- 5 years) HIV-1 DNA copies/million PBMCs, respectively.

Consequently, the total HIV-1 DNA copies/million PBMCs, increased in between the first (Figure 17 A) and last (Figure 17 B) time points for the group that achieved VS < 1 year of age with medians of 18.8 and 57.2 respectively. There was a decrease of total HIV-1 DNA copies/million PBMCs between first (Figure 17 A) and last (Figure 17 B) time points for the group that achieved VS between 1-5 years with medians of 193.0 and 178.9 HIV-1 DNA copies/million PBMCs, respectively.

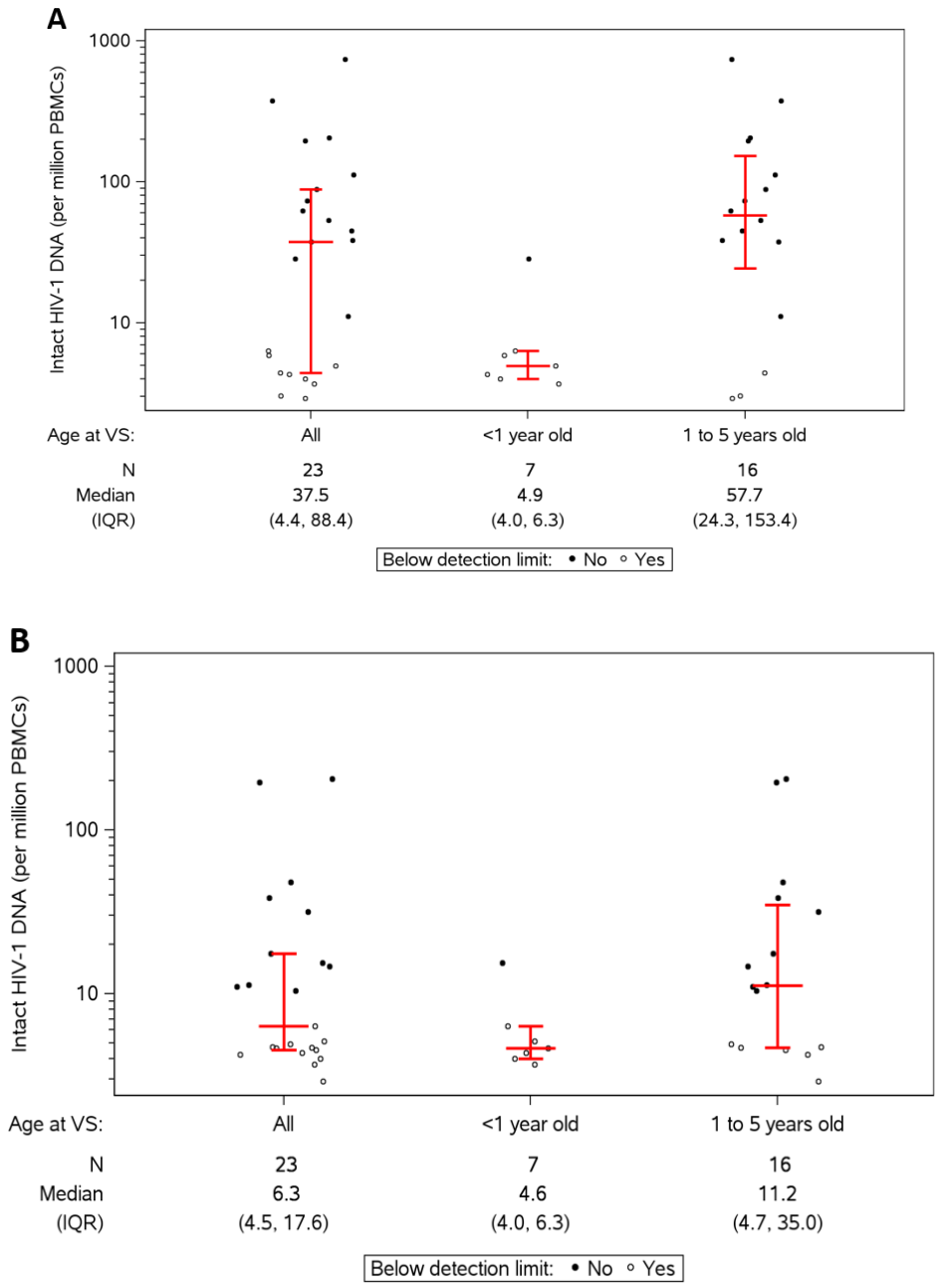


Figure 14. Distribution of PBMC-associated intact HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant).

Panel A: Data for first chronological point and Panel B: Data for last chronological point. The plot is based on age at virologic suppression at the first and last chronological samples collection. Certain participants did not have 2 time points therefore may not be accurately depicted here. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range

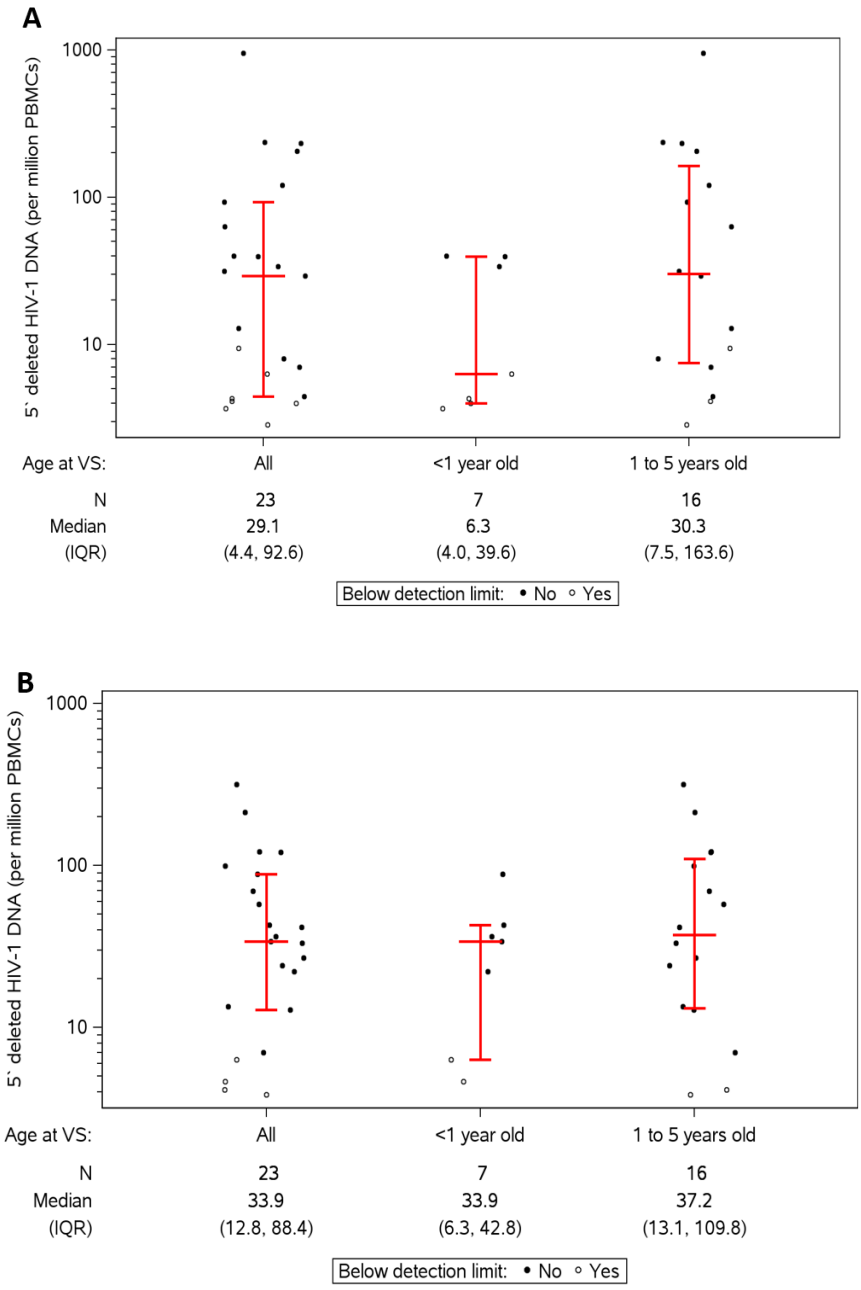


Figure 15. Distribution of PBMC- associated 5'-deleted (defective) HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant).

Panel A: Data for first chronological point and Panel B: Data for last chronological point. The plot is based on age at virologic suppression at the first and last chronological samples collection. Certain participants did not have 2 time points therefore may not be accurately depicted here. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range

Next, the data was stratified and analyzed by duration of virologic suppression (VS) and was divided into three groups <2 years of VS (22), 2-5 years of VS (34), >5 years of VS (6).

When stratified by the duration of VS, the distributions of the different species followed different trends with increasing duration of virologic suppression. The intact HIV-1 DNA demonstrated a decline between each group of duration of VS with medians of 41.2 (VS<2 years), 8.3 (VS=2-5 years) and 4.9 (VS>5 years) intact HIV-1 DNA copies/million PBMCs (Figure 18 A). The decline in intact HIV-1 DNA copies/million PMBCs between VS<2 years and VS 2- 5 years was statistically significant with a p value of 0.0012 ($\alpha = 0.05$). The decline between intact HIV -1 DNA copies/million PMBCs between VS<2 years and VS>5 years was statistically significant with a p value of 0.0236 ($\alpha = 0.05$) while the decline between VS 2-5 years and VS>5 years was not statistically significant.

A similar trend was also noted for 5' deleted HIV-1 DNA copies/million PBMCs (Figure 18 D). The total and 3' deleted/hypermutated HIV-1 DNA copies/million PBMCs showed an initial decline followed by an increase in their concentrations (Figure 18 B and C).

The total HIV-1 DNA copies/million PBMCs declined between the VS<2 years and VS for 2-5 years with medians of 138.5 and 61.5 HIV-1 DNA copies/million PBMCs, respectively. Then an increase was noted between VS for 2-5 years and VS>5 years with medians of 61.5 and 93.9 HIV-1 DNA copies/million PBMCs, respectively (Figure 18 B). The changes amongst the three subgroups

of duration of virologic suppression for total HIV-1 DNA were not statistically significant.

The 3’ deleted/hypermutated HIV-1 DNA copies/million PBMCs showed an initial decline from VS<2 years to VS = 2-5 years with medians of 57.4 and 29.4 HIV-1 DNA copies/million PBMCs, respectively. This was followed by an increase in 3’ deleted/hypermutated HIV-1 DNA copies/million PBMCs between VS for 2-5 years and VS> 5 years with medians of 29.4 and 61.2 HIV-1 DNA copies/million PBMCs, respectively (Figure 18 C). The changes in 3’deleted/hypermutated proviruses amongst the three subgroups of duration of virologic suppression were also not statistically significant.

The 5’ deleted HIV-1 DNA copies/million PBMCs showed a relatively slow decline with increasing duration of suppression with medians of 35.5 (VS<2 years), 29.4 (VS = 2-5 years) and 26.8 (VS>5 years) HIV-1 DNA copies/million PBMCs (Figure 18 D). The changes in 5’deleted copies/million PMBCs amongst the three subgroups of duration of virologic suppression were not statistically significant.

Duration of VS	Median Intact HIV-1 DNA CPM	Median 5' deleted HIV-1 DNA CPM	Median 3' deleted HIV-1 DNA CPM	Median Total HIV-1 DNA CPM
less than 2 years	41.2	35.5	57.4	138.5
2 - 5 years	8.3	29.4	29.4	61.5
greater than 5 years	4.9	26.8	61.8	93.9

Table 10: Medians of HIV-1 DNA species segregated based on duration of virologic suppression.

CPM = copies/million PBMCs

When the proportion of total HIV-1 DNA that is intact was analyzed between the three groups of duration of VS, a decline in the proportion of intact HIV-1 DNA with increasing duration of VS was noted (Figure 19). However, not every participant had a sample for each group and therefore these results could potentially be skewed by this as well as the small cohort size.

The %intact values that were at or above 100% were a result of the total (sum of intact + 3’ deleted/hypermutated + 5’ deleted HIV-1 DNA copies/million PBMCs) HIV-1 DNA copies/million PMBCs being equal to or less than the intact HIV-1 DNA copies/million PMBCs. This can be attributed to the poor/lack of detection of 3’ deleted/hypermutated or 5’ deleted or both 5’ deleted and 3’ deleted/hypermutated HIV-1 DNA copies/million PBMCs.

The IPDA was designed to quantify intact proviruses when the reservoir is stable, and the proportion of intact proviruses is not very high. However, in the current study, it could be possible that the time points analysed were around the time that the reservoir was still stabilizing and had not recached the second stage of the biphasic decay leading to the presence of many copies of intact proviruses. As a result of this the quantification by the IPDA might have become less accurate.

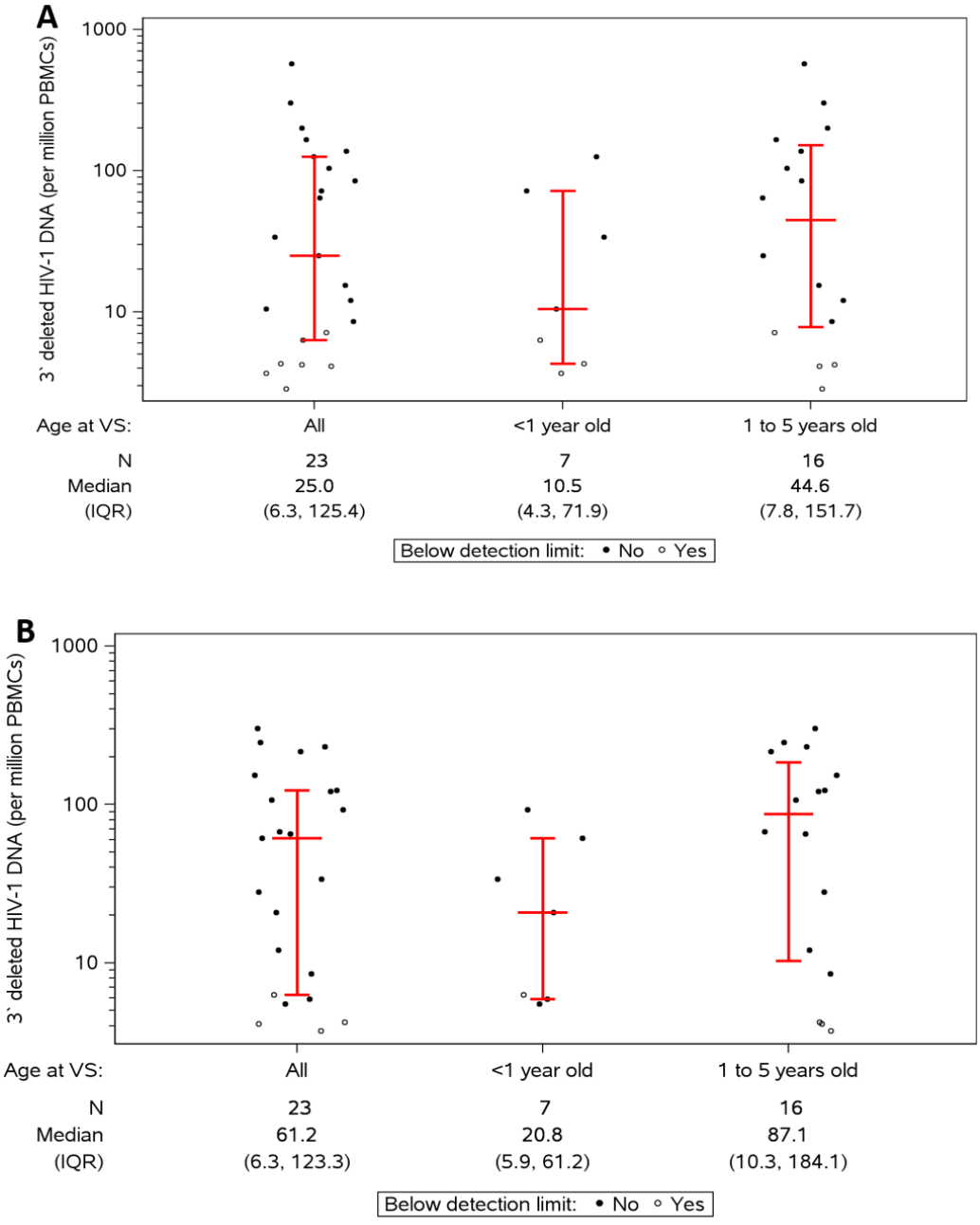


Figure 16. Distribution of PBMC-associated 3'-deleted/hypermutated (defective) HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant).

Panel A: Data for first chronological point and Panel B: Data for last chronological point. The plot is based on age at virologic suppression at the first and last chronological samples collection. Certain participants did not have 2 time points therefore may not be accurately depicted here. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range

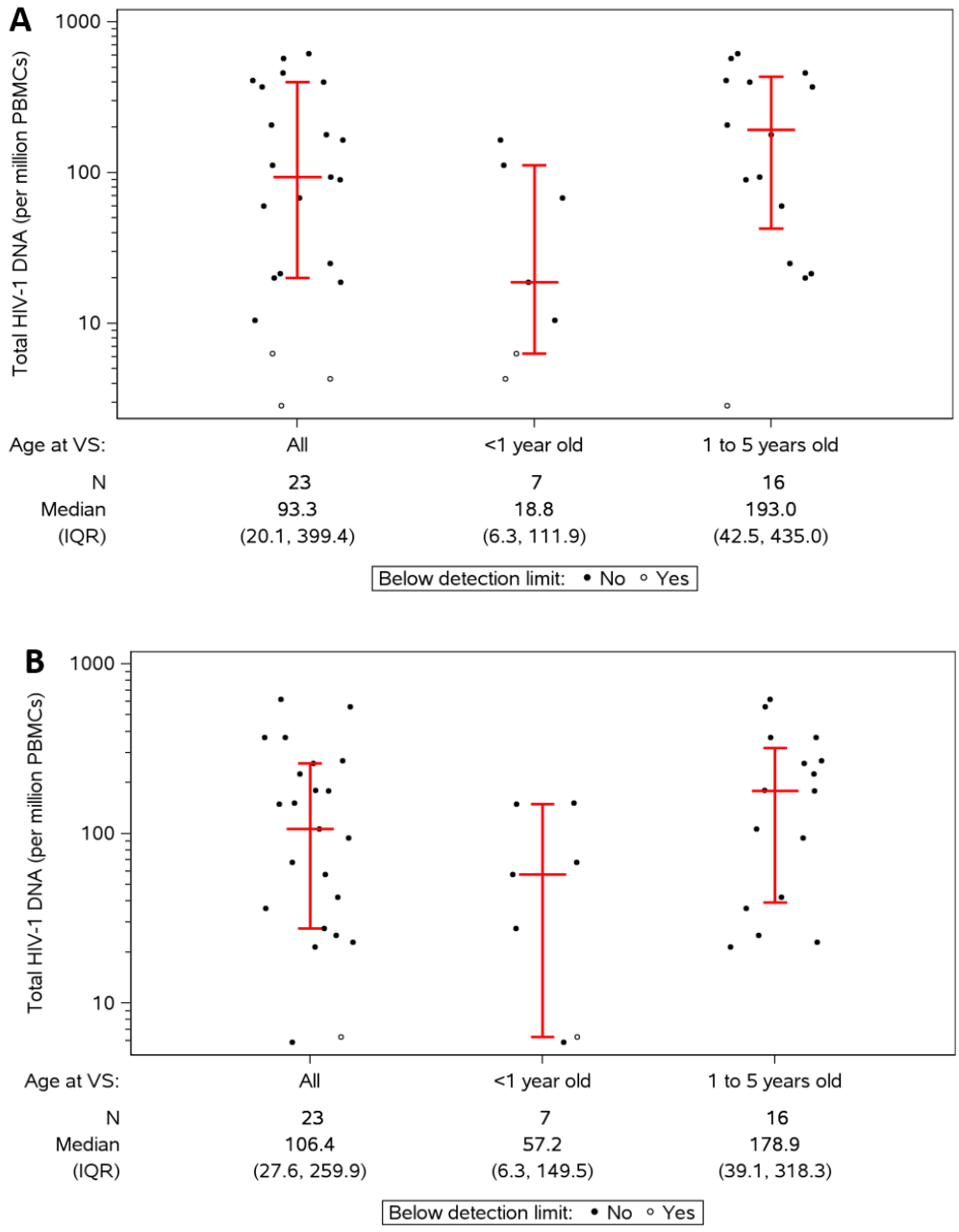


Figure 17. Distribution of PBMC- associated total (intact + defective) HIV-1 DNA, by age at virologic suppression (first and last chronologic sample per participant).

Panel A: Data for first chronological point and Panel B: Data for last chronological point. The plot is based on age at virologic suppression at the first and last chronological samples collection. Certain participants did not have 2 time points therefore may not be accurately depicted here. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range

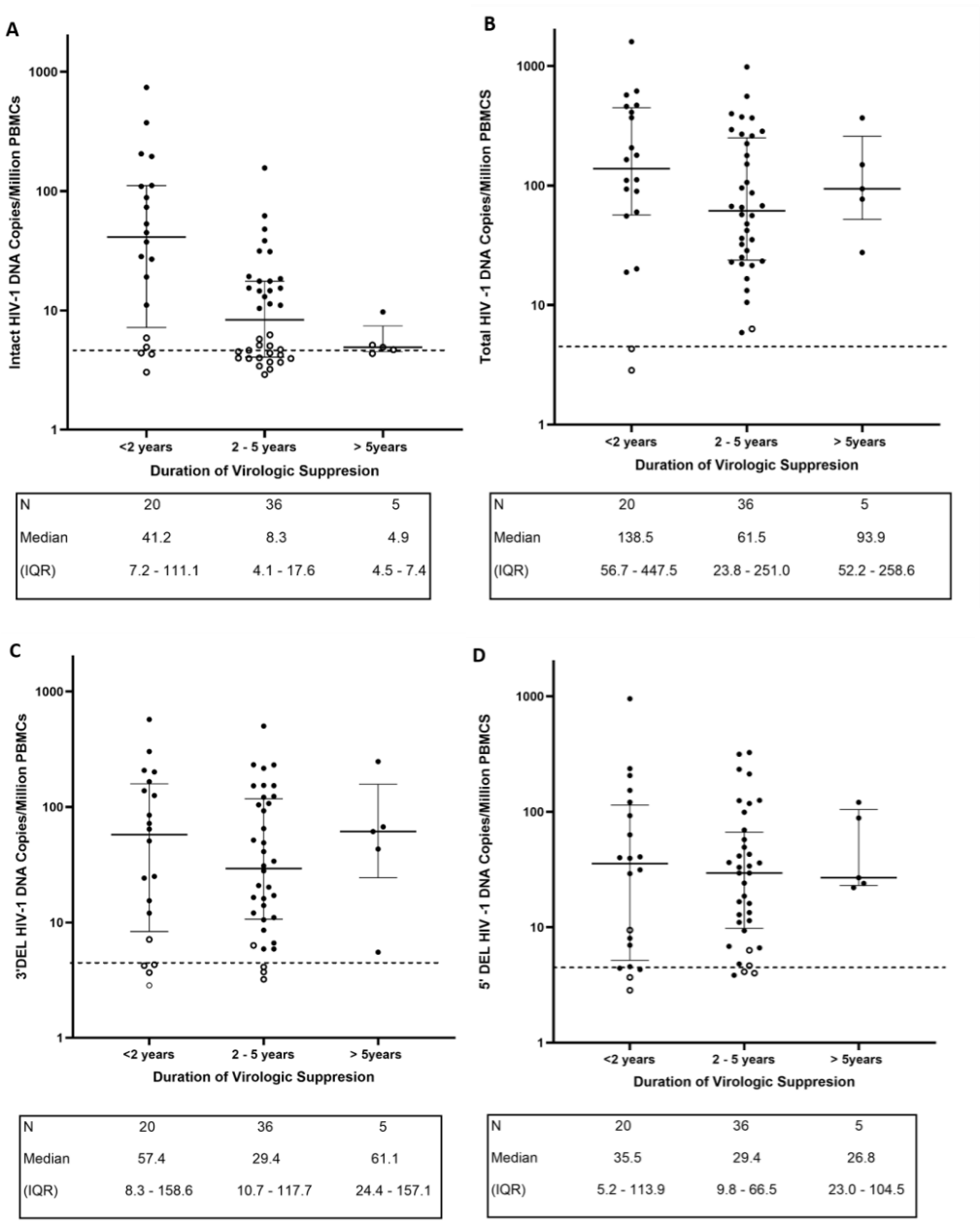


Figure 18. Distribution of various HIV-1 DNA species as a function of duration of Virologic Suppression.

Panel A: Intact HIV-1 DNA copies/million PBMCs as a function of duration of suppression. Panel B: Total (intact + defective) HIV-1 DNA copies/million PBMCs as a function of duration of suppression. Panel C: 5' deleted (defective) HIV-1 DNA copies/million PBMCs as a function of duration of suppression. Panel D: 3' deleted/hypermutated (defective) HIV-1 DNA copies/million PBMCs as a function of duration of suppression. Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The dotted lines depict the average limit of detection for the cohort. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. IQR = Interquartile range.

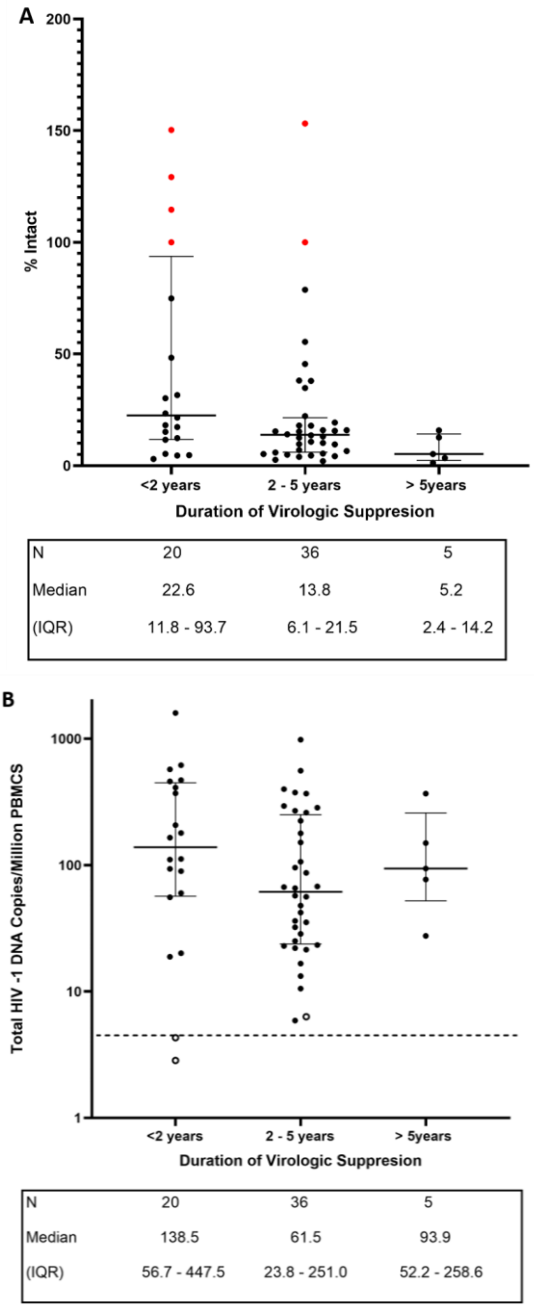


Figure 19. Proportion of total HIV-1 DNA that was intact.

Panel A: % Intact HIV-1 DNA as a function of duration of suppression. Panel B: Total (intact + defective) HIV-1 DNA copies/million PBMCs as a function of duration of suppression Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The dotted lines depict the average limit of detection for the cohort. The open symbols indicate samples in which HIV-1 DNA was not detected. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. The red dots depict data points where the intact was greater than the total. There was an outlier, a time point where the %intact was 13155 and was excluded from the graph. However, this outlier did not impact the median values. IQR = Interquartile range. CPM = Copies/Million

2.4 Individual Data

Based on the individual plots (Supplementary figures 1-5) the reservoir dynamics for each participant were quite variable. However, most data (14/25) could be grouped based on similar trends. The three major trends observed were: variable total HIV-1 DNA copies/million PBMCs (CPM PBMCs) over time (supplementary figure 1), declining HIV-1 DNA CPM PBMCs over time (supplementary figure 2) and stable HIV-1 DNA CPM PBMCs over time (supplementary figure 3). The total HIV-1 DNA CPM PBMCs is a measure of the sum of intact HIV-1 DNA copies/million PBMCs, 3' deleted/hypermutated HIV-1 DNA copies/million PBMCs and 5'deleted HIV-1 DNA copies/million PBMCs. Therefore, changes in any of these three species would be reflected in the total HIV-1 DNA CPM PBMCs. In the participants (6/25) that demonstrated variable total HIV-1 DNA CPM PBMCs over time, it was observed that the total HIV-1 DNA was in a state of constant flux with an increase followed by decrease in the HIV-1 intact and defective species or vice versa. The participants (4/25) that demonstrated declining total HIV-1 DNA CPM PBMCs over time had declining concentrations of all species which is what one would expect with increasing time on ART and duration of VS.

With respect to stable HIV-1 DNA CPM PBMCs (4/25), it was noted that the proportion of the intact to defective (3' deleted/hypermutated or 5' deleted) HIV-1 DNA copies/million PBMCs was such that if a decline were observed in one species, a proportional change would be observed in the other two species to compensate for the change in intact proviruses. In this case, a decline in the intact HIV-1 DNA CPM PBMCs was compensated by changes in the

proportions of the 3' deleted/hypermutated and 5' deleted HIV-1 DNA CPM PBMCs.

The remaining participants (11/25) were stratified into three groups: those that only had one time point (4/25, Supplementary figure 4), those that needed further testing (5/25, Supplementary figure 5) and those that had purported primer-probe binding issues (2/25). For the group with only one time point no major conclusions could be derived from their data. The two participants of the 25 that demonstrated what appeared to be primer-probe binding issues were tested further to confirm this[171].

3. IPDA vs Pol LTR on PHACS samples

As the IPDA is a relatively new ddPCR assay that is used for studying the HIV-1 reservoir, it was interesting to see how its results would correlate with those of previous assays used for studying the HIV-1 reservoir. Since there was data from POL-LTR assay (previous study) used to study the reservoir on the PHACS cohort[60], correlations were run between the data from the POL-LTR assay and the IPDA. The POL-LTR assay gave a measure of the total HIV-1 DNA copies/million PBMCs vs the IPDA which gave a measure of total, intact, 3' deleted/hypermutated and 5' deleted HIV-1 DNA copies/million PBMCs. For each of these species a non-parametric Spearman's rank test was performed to note the correlation and a unpaired t-test was performed to determine the p-value with $\alpha = 0.05$.

The total HIV-1 DNA copies/million PBMCs from the IPDA when compared to the total HIV-1 DNA copies/million PBMCs from the POL-LTR assay showed a positive Spearman's correlation co-efficient (r) = 0.8259 , p<0.0001 (Figure 20 A). When the intact HIV-1 DNA copies/million PBMCs were compared with the total HIV-1 DNA copies/million PBMCs from the POL LTR assay, a positive correlation was noted with r = 0.6981 and p<0.0001 (Figure 20 B). The 5' deleted (Figure 20 C) and 3' deleted/hypermutated (Figure 20 D) HIV-1 DNA copies/million PBMCs when compared with the total HIV-1 DNA copies/million PBMCs from the POL-LTR assay, also showed a positive correlation; however, they had lower r values (compared to intact HIV-1 DNA species) of 0.6779 and 0.6186, p<0.0001, respectively.

4. Determination of primer-probe mismatch

During analysis, a 2D plot shows each species in a different XY quadrant based on the PCR reaction and the fluorophore. Quadrant 1 shows the 3' deleted and hypermutated proviruses, quadrant 2 shows intact HIV-1 proviruses, quadrant 3 shows some proviruses that did not amplify at both regions + negative reactions for HIV-1 DNA and quadrant 4 shows 5' deleted proviruses. (Figure 4). If the primer/probe is unable to bind to its respective region in the gag/env portion of the HIV-1 genome, fluorescence will not be observed in that DNA species' quadrant as well as quadrant 2. This would indicate a primer-probe binding issue as was observed for two participants[171].

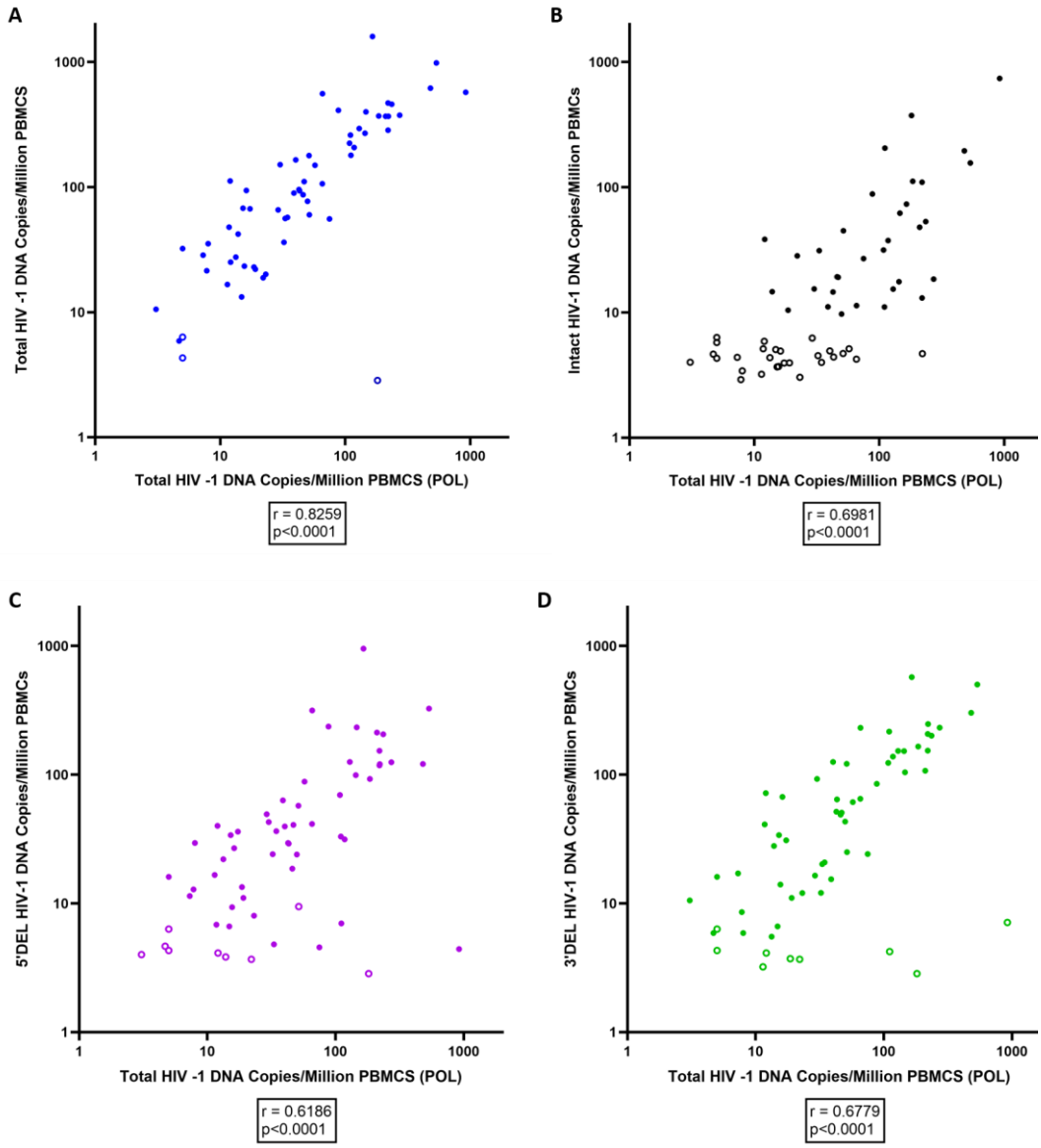


Figure 20. Correlations between Total HIV-1 DNA Copies/Million PBMCs detected by POL-LTR assay vs HIV-1 DNA Copies/Million PBMCs of all species by IPDA.

Panel A: Total HIV-1 DNA Copies/Million PBMCs (IPDA) vs Total HIV-1 DNA Copies/Million PBMCs (POL) Panel B: Intact HIV-1 DNA Copies/Million PBMCs vs Total HIV-1 DNA Copies/Million PBMCs (POL) Panel C: 5' deleted HIV-1 DNA Copies/Million PBMCs vs Total HIV-1 DNA Copies/Million PBMCs (POL) Panel D: 3' deleted/hypermutated HIV-1 DNA Copies/Million PBMCs vs Total HIV-1 DNA Copies/Million PBMCs (POL). Samples with 2LTR values ≥10 and participants with purported primer-probe binding issues were excluded. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. Assay results below limit of detection were set to their respective detection limits which vary with the number of cells analyzed. Non-parametric Spearman test was run on all correlations to determine the Spearman correlation co-efficient and two-tailed t Test was used to determine the p-value with $\alpha = 0.05$. CPM = Copies/Million

The 2D amplitude plot for two participants, PHACS sample 3 and PHACS sample 13 did not show droplets in the quadrant for intact HIV-1 DNA (Q2) and env HIV-1 DNA (Q4), and intact HIV-1 DNA (Q2) and gag HIV-1 DNA (Q1), respectively, as seen in Figure 21. Based on a previous study by Simonetti et al., it appeared to be a primer-probe mismatch[171]. To confirm this, the samples were tested further. Two PCRs were performed using a set of primers that covered more area than that bound by the IPDA primers, and near full-length primers respectively. Primer-probe mismatches were confirmed for both the participants based on these PCR results. PHACS sample 25 which did not show any droplets in all three quadrants, Q1, Q2 and Q4 was also tested for primer-probe mismatch.

4.1 Amplifying the primer-probe binding region

To ensure that there was enough remnant genomic DNA for further analyses, whole genome amplification was performed followed by the outer and nested PCRs using the primers shown in Table 4. The nested PCR products were run on a 1% gel as seen in Figure 22. All the time points for PHACS sample 3 showed bands for gag, two of four time points for PHACS sample 13 showed gag bands at the right position of 1183 bp. Only one time point of PHACS sample 13 showed a band for env at 957 bp. PHACS sample 25 did not show any amplicons and was not tested further. On repeating the nested PCR with a 1:10 diluted outer PCR product, the env PCR for PHACS sample 3 still failed to show amplicon bands after the gel electrophoresis as did the u5gag PCR for the last two time points for PHACS sample13 (Figure 23).

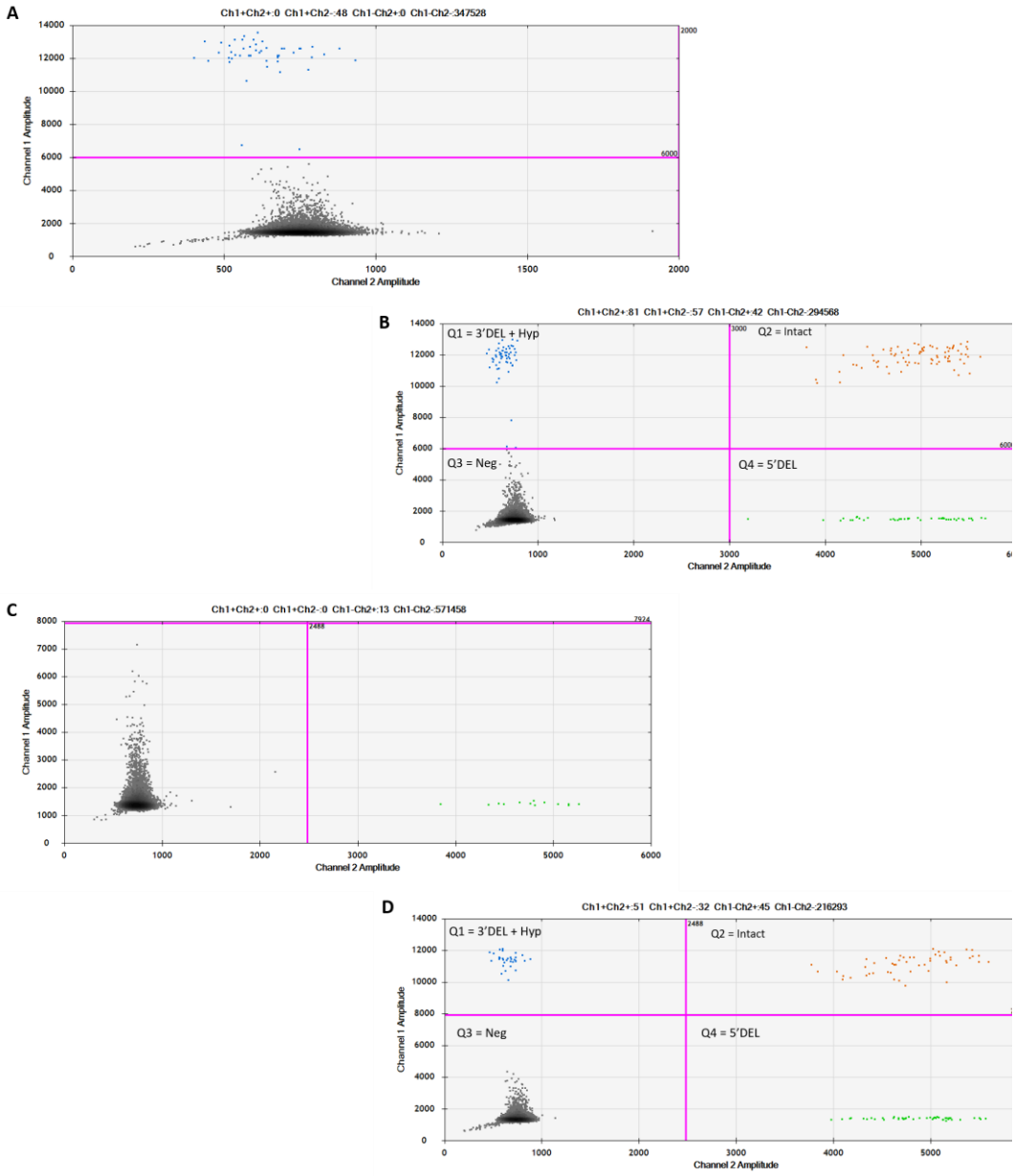


Figure 21. Primer-Probe mismatch in participant samples.

Panel A: 2D amplitude plot showing potential primer probe mismatch for ENV region Panel B: 2D amplitude plot for controls of the same run. Panel C: 2D amplitude plot showing potential primer probe mismatch for u5GAG region Panel D: 2D amplitude plot for controls of the same run. The lack of droplets in Q3 + Q2 and Q1 + Q3 respectively indicate that there may be a primer-probe mismatch involved.

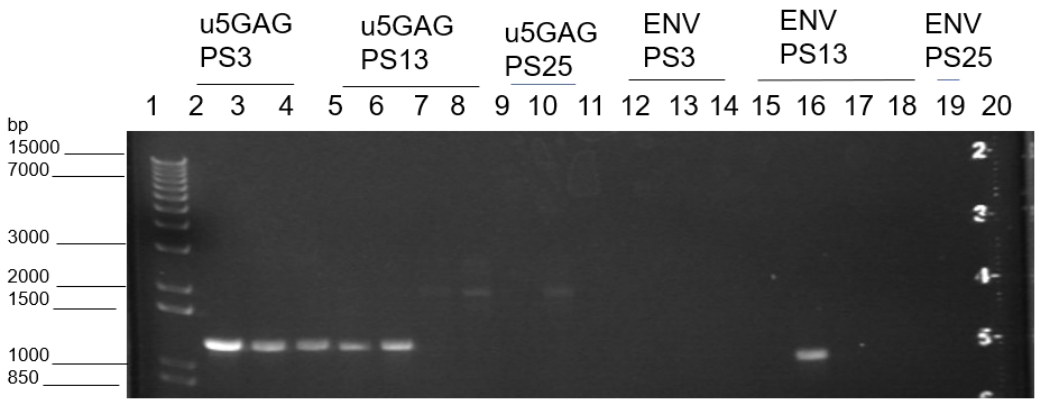


Figure 22. Gel Electrophoresis of Nested PCR for u5gag and env regions.

Lane 1: Ladder, Lane 2-4: u5GAG PS3, Lane 5-8:u5GAG PS13, Lane 9-10:u5GAG PS25, Lane 12-14: ENV PS3, Lane 15-18:ENV PS13, Lane19: ENV PS25

In the nested u5GAG PCR, all the time points for PS3 had a band for gag as expected. The first 2 time point for PS13 showed amplicon bands at the correct size of 1183bp while the other two time points had no band and a light band at the wrong size respectively. PS 25 did not show an amplicon band for u5GAG.

In the nested Env PCR, most of the time points except one time point in PS 13 did not show an amplicon band. The PS 13 env amplicon was at the correct size of 957 bp. The nested PCR was then repeated. PS = PHACS sample

Based on these results, we thought that there may have been a mismatch even with the sequencing primers. Therefore, a near full-length four fragment PCR was performed to help amplify the proviruses that could not be amplified using the sequencing primers in Table 4. Gel electrophoresis was performed on the nested 4 fragment PCR products (Figure 24). PHACS sample 13 did not show amplicons for Fragment A or B (Figure 3). PHACS sample 3 showed amplicons for Fragment C and D (Figure 3) at 6385bp and 4778bp respectively, as well as the alternative env primers at 2841bp. This confirmed that there was a primer-probe mismatch for PHACS sample 3.

As for PHACS sample 13, the two time points that showed gag amplicons in the first set of PCRs (using the sequencing primers in table 4) were sent for Sanger sequencing while it was believed that the last two time points either had no gag regions or another provirus may have been amplified by the IPDA.

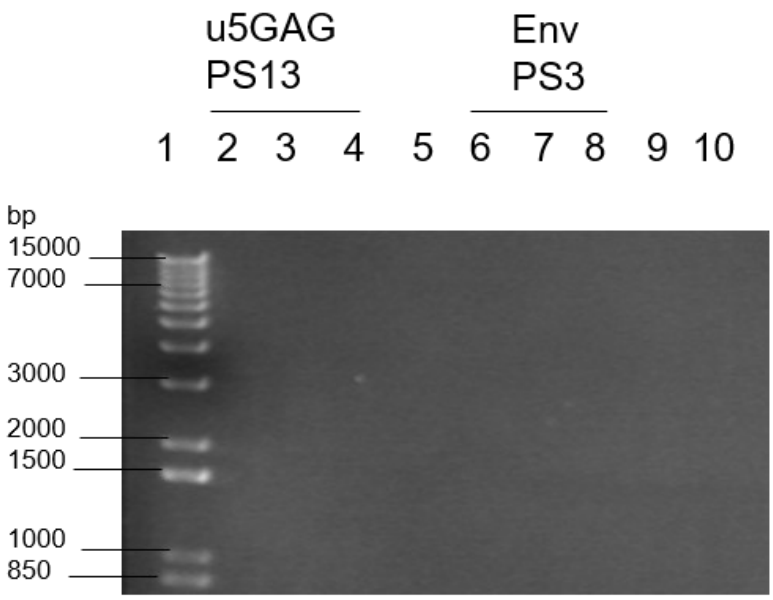


Figure 23. Gel Electrophoresis of Repeated Nested PCR for u5gag and env regions.

Lane 1: Ladder, Lane 2-3: u5GAG PS13, Lane 4: control, Lane 6-8: Env PS3, Lane 9; control, Lane 12-14: ENV PS3, Lane 15-18:ENV PS13, Lane19: ENV PS25

The nested env PCR was repeated with a 1:10 dilution for all time points of PS3 and the u5gag nested PCR was repeated with a 1:10 dilution for 2 time points of PS13. Based on the gel, it was noted that the env PCR had failed while PS13 timepoints failed to show u5gag. PS= PHACS sample

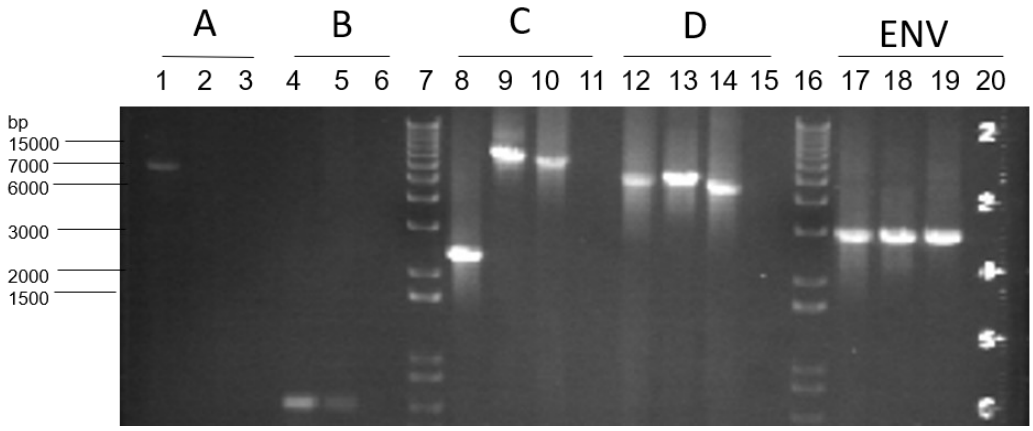


Figure 24. Near full-length four fragment PCR Gel Electrophoresis.

Lane 1-2 Fragment A, Lane 3 Fragment A control; Lane 4-5 Fragment B, Lane 6: Fragment B control Lane 7: Ladder; Lane 8-10 Fragment C; Lane 11 Fragment C control; Lane 12-14 Fragment D, Lane 15 control; Lane 16 Ladder; Lane 17-19 Alternative env primers; Lane 20 Control for env primers.

All PS3 samples showed a positive band for Fragment C at 6385 bp although the band in lane 8 was truncated; Fragment D at 4778 bp and the ENV primers at 2841bp. PS13 bands were not observed for Fragment A and Fragment B at the appropriate band size. PS: PHACS sample

4.2. Sanger sequencing

Sanger sequencing was performed on PHACS sample 3 and PHACS sample 13 using the nested primer shown in Table 4. The sequences were then aligned with the reference HXB2 and the primer-probe of each set. For PHACS sample 3, it was noted that there was a primer mismatch (forward primer) as well as a probe mismatch which led to the lack of detection of the env region by the IPDA. In the forward primer mismatch (Figure 25 A), within the highlighted region, a consistent single nucleotide change from G to A was observed for all time points. For PHACS sample 3 W78 another G to A change was observed in the primer binding region. The probe mismatch (Figure 25 B) showed three single

nucleotide polymorphisms of pyrimidine to purine changes, with two C to G and one C to A mutation within the highlighted region.

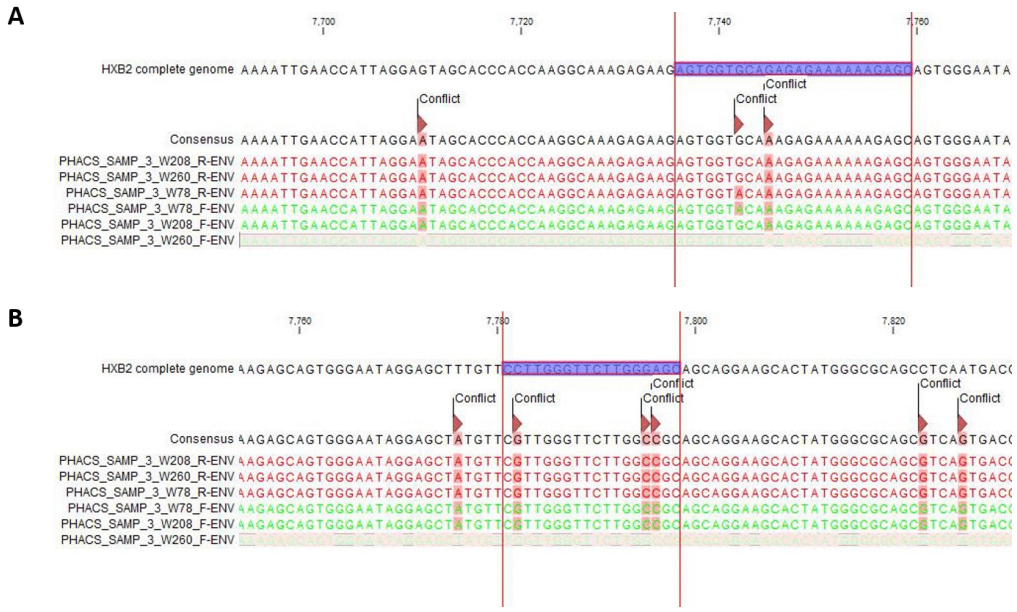


Figure 25. Sanger Sequencing and alignment results for env region mismatch.

Panel A: Sequence alignment of forward and reverse amplicons vs HXB2 genome as a reference and the env forward primer. The highlighted region depicts the primer binding region. All the sample sequences showed a single nucleotide polymorphism at the same site G to A. The sample at time point W78 had another mismatch with a G to A mutation.

Panel B: Sequence alignment of forward and reverse amplicons vs HXB2 genome as a reference and the env probe. The highlighted region depicts the probe binding region. All the samples showed three single nucleotide polymorphisms of pyrimidine to purine with two C to G and one C to A mutation.

For PHACS sample 13, two time points were sequenced and while there was no primer mismatch, it was noted that there was a probe mismatch which led to lack of detection of the gag region by the IPDA (Figure 26). In the case of PHACS sample 13, only the forward primer sequences could be aligned as the probe bound closer to that region and the reverse primer amplicon was not long enough to reach the probe binding region.

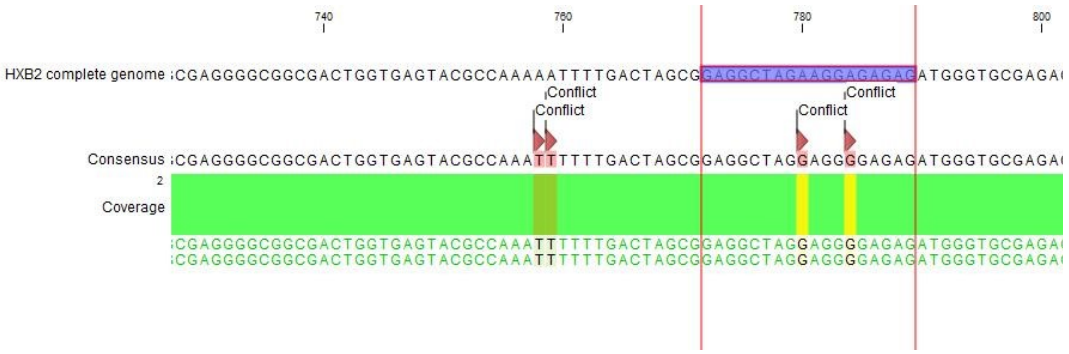


Figure 26. Sanger Sequencing and alignment results for u5gag region mismatch.

The sequence of the forward amplicon was aligned against the HXB2 genome as a reference and the probe of the u5GAG. Two mismatches close to the 5' end of the probe binding region seemed to have contributed to primer-probe mismatch. Reverse amplicon is not depicted as it did not cover the probe binding region. No mismatches were seen in the primer binding regions. Sample sequences were consistent across the 2 time points tested.

The next step will be to get the individual primer probes developed for each of the participants and the IPDA will be re-run.

IV. Discussion

Optimization of the IPDA

To optimize the Intact Proviral DNA Assay (IPDA), it was crucial to determine: a) the best control for the assay; and b) the performance characteristics of the assay. The JLAT cell line was chosen to develop controls as previously published data have shown that JLAT cells have only one copy of integrated HIV-1 DNA per cell and are stable over many passages[225]. Plasmids were not used as a control because the plasmid vials contain very concentrated plasmid levels (high copy number) which can result in cross contamination with controls for other assays in the laboratory. Additionally, the JLAT cells were preferred over the use of plasmids, as the JLAT cells have an integrated copy of HIV-1 DNA which mimics the reservoir as seen in-vivo[224]. Thus, JLAT cells were chosen to develop controls for the IPDA.

Once the controls were developed, the performance characteristics were determined. The IPDA was run using different amounts of DNA input. The genomic DNA used for the assay was isolated from the JLAT standards with varying concentrations of proviral DNA copies/million PBMCs. The amounts of DNA input were 1000ng, 500ng and 200ng. The assay was able to detect HIV-1 DNA copies for all three DNA input amounts (1000ng, 500ng, 200ng). Regardless of the amount of input DNA, the expected vs observed HIV-1 DNA copies/million PBMCs fell within the acceptable range (± 10% error).

The linearity was maintained for all three input DNA amounts which was very important and proved that the dilution scheme worked for all three amounts of DNA input.

In the case of 1000ng DNA input, the assay could not reliably detect intact HIV-1 DNA copies below a concentration of 2.5 intact HIV-1 DNA copies/million PBMCs. For the 500ng and 200ng DNA input, the assay detection was less predictable with some runs showing detection below 5 intact HIV-1 DNA copies/million PBMCs. This suggests that: 1) the assay is unable to reliably detect below 10 HIV-1 DNA copies/million PBMCs; and 2) the assay has a limit of detection between 5 - 2.5 intact HIV-1 DNA copies/million depending on the amount of genomic DNA input. Similar limits of detection were observed in the Siliciano laboratory, where the assay was developed[69].

An important point to note is that perinatally infected children (early suppressed) have low proviral loads, usually below 25 HIV-1 DNA copies/million PBMCs as seen later with PHACS samples, therefore it is important for the IPDA to reliably detect HIV-1 DNA below 10 HIV-1 DNA copies/million PBMCs. One of the reasons that could explain why the detection becomes more variable below 10 HIV-1 DNA copies/million PBMCs is that the IPDA was developed for use during chronic treatment. It has been noted that hypermutations and deletions rates may be different in early treated participants, which could affect the performance of the IPDA[168].

Based on current observations, the assay may not be well suited as a standalone assay for detecting proviral loads in early suppressed participants; however, it is well suited for studying the reservoir in later suppressed

participants. As mentioned earlier, this issue may potentially be resolved by running experiments with a greater number of replicates.

As stated before, the maintenance of linearity is important when the assay is applied to participant samples that may not have sufficient DNA to input 1000ng per well for multiple replicates. Knowing that the assay worked well even at lower input DNA amounts allows the assay to be applied to samples regardless of DNA input to study the reservoir, albeit with limitations for samples containing lower than five intact proviruses.

The next sets of performance characteristics that affected the readout of the assay were shearing of DNA and number of cells analyzed. Based on the limit of detection of the ddPCR assays optimized in our laboratory, it was empirically determined that to get a good estimate of the composition of the reservoir, i.e., the intact proviruses that can re-establish viremia when ART is stopped, ideally 800,000 to 1,000,000 cellular equivalents (from PBMCs or CD4 T cells) of genomic DNA should be analyzed. But it is not always possible to analyze 1 million cellular equivalents, as it requires genomic DNA extraction from 3-5 million PBMCs.

In pediatric HIV-1 infections studying the reservoir is complex and challenging not only due to the evolving immunology from neonates to young adults or routes of transmission but also owing to blood sampling. Neonates can be infected by an HIV-1 positive mother either in-utero or during the intra-partum period (perinatal infection) or through breast feeding[10]. If the infection is left untreated the neonate dies within 1-3 years of birth[7], to prevent this, ART is started as soon as the infection is diagnosed, which can be anywhere between

1-3 months of birth or if the neonate is at high risk for infection, within days/weeks of birth[3,92,227,228]. The consequences of this early treatment are very low proviral loads[7,216].

With respect to blood sampling, based on the weight of the neonate, about 2.5-5mL of blood can be drawn from a neonate. This volume increases as the weight of the infant increases but is usually not more than 50mL in adolescents[229-231]. The low volumes make it difficult to extract 5 million cells for analysis of the reservoir in perinatal infections and the output from processing the blood sample has to be maximized. In some cases, less than 5 million cells are obtained. During optimization of the IPDA in the present study, standards with 3 million cells were used to determine if the assay would function appropriately. The number of cells analyzed fell between 150,000 - 350,000 which made it challenging to study the proviral reservoir; however even though the number of cells were not ideal, we were still able to study the reservoir. This proved that although cells analyzed are an important factor, in an unideal situation where the pellet size is less than 5 million cells, the reservoir can still be studied to inform ART-free remission strategies, and the number of cells analyzed need not be a strong limiting factor.

The shearing index of DNA is extremely important in the IPDA. DNA is unsheared in a cell but when it is processed for isolation and ddPCR, the DNA breaks into fragments. This shearing, if not accounted for results in an underestimation of the number of intact HIV-1 DNA copies present in the sample, and the size of the reservoir. Hence it is important to account for shearing accurately and it is done by using the housekeeping gene RNase P30

as described earlier[69]. Accounting for shearing is especially important in pediatric HIV-1 infections owing to early treatment resulting in low proviral loads. Therefore, it is very important to minimize shearing from the first step of processing the cells i.e., DNA isolation, which we can improve in future studies with use of the non-column-based approaches for genomic DNA extraction from cells.

We showed that when the Qiagen Qi-Amp kit was used for the isolation of genomic DNA from the standards, a median %unsheared of 61.62% was obtained. This meant that about 39% of the DNA had been fragmented. This rate of shearing can be attributed to mechanical stresses, freeze thaw cycles and to the passing of DNA through the silica column used in the Qiagen Qi-amp Midi kit. Upon isolating the DNA using the Qiagen Gentra Puregene core B kit, the median of %unsheared was 75.55%. This could be because the Gentra Puregene kit reduces the mechanical stresses, minimizes vortexing and eliminates the use of silicon columns. The Gentra Puregene kit was tested by three different performers and we saw similar findings of reduced shearing. Hence the comparison between the two kits used for isolation demonstrated that the lack of columns can affect the %unsheared.

However, when the IPDA was established by Bruner et al., it was noted that the Qiagen Qi-amp kit was used for isolation of genomic DNA (samples from adults) and the shearing was accounted for using the RNase P30 gene PCR only and not at the step of DNA isolation[69]. This was less impactful on the results because the adult participants had very high levels of intact proviruses (a

median of 100 intact HIV- 1 DNA copies/million CD4 T cells), which was not the case for pediatrics as observed upon applying the IPDA to the PHACS samples.

When the IPDA was applied to the PHACS cohort study, the shearing of the DNA was quite high with the majority of samples falling within 45-60% sheared DNA. This was expected as these samples had been processed and isolated over five years[60]. The DNA underwent several freeze thaw cycles, and other mechanical stresses as different assays were performed on this sample set.

The observable impact of this shearing may have influenced the lack of detection of intact HIV-1 proviruses in the cohort with only 54.1% of the 61 samples showing detectable intact HIV-1 DNA copies/million PBMCs. As mentioned earlier, the low detection of intact HIV-1 DNA copies would be expected with such high shearing levels since proviral loads in children are very low. Therefore, when applying an assay like the IPDA to study the reservoir size and dynamics, it is best to minimize shearing from step 1 (i.e., DNA isolation) and to process a fresh aliquot of the samples when possible, to avoid underestimating the reservoir size.

The Gentra Puregene kit reduces the amount of shearing at the step of DNA isolation, which reduces the need to rely on shear correction done post ddPCR, while analyzing the samples. This is important because if the DNA has a very low %unsheared as seen for PHACS sample 10 at one time point (data not shown), the shear correction overcompensates and biases the results towards intact HIV-1 DNA copies which can affect the results of the study. However, this does not imply that using the Qi-amp kit instead of the Gentra Puregene kit is not acceptable, since we can still study the reservoir.

It just means that in pediatric infections, it is preferred to account for shearing from step 1 which can be done by using a kit like the Gentra Puregene.

Applying the IPDA to children living with perinatal HIV-1 in the PHACS-AMP cohort

In the past studies on perinatal HIV-1 infections across different cohorts, durations of suppression and subtypes of HIV-1, it was observed that the reservoir size decreased with early treatment and suppression vs deferred treatment or late suppression[55,59,191,196,199,201-203,207,215,216,220,222].

However, the reason for decrease in the size of the reservoir was not well elucidated hence to understand what was going on, in a seminal study, Uprety et al., studied the decay dynamics of the reservoir in long-term suppressed children from the Pediatric HIV AIDS Cohort Study[60]. They observed that within the first two years of virologic suppression (VS), the HIV-1 DNA decayed much faster in early suppressed group than it did in the later suppressed group[60]. This explained the reason for the smaller reservoir sizes in early treated children but the exact reason for the decay of total HIV-1 DNA was not clear. Therefore, to further elucidate the differential rate of decay, IPDA was performed on the PHACS group.

The Pediatric HIV AIDS cohort study, Adolescent Master Protocol is a long-term longitudinal study of children and adolescents living with perinatal HIV-1 in the U.S.[210] As mentioned earlier in the previous studies done on the PHACS cohort, the investigators were able to ascertain that early treatment reduced the

proviral size. The previous study done in our laboratory by Uprety et al., used the banked samples from this cohort to study the decay dynamics of the proviral DNA based on age at virologic suppression (VS) in order to understand the low proviral loads seen in children who achieved virologic control from infancy[60].

The previous study utilized a single-plex ddPCR, total HIV-1 DNA assay: POL-LTR to measure proviral loads over years of virologic suppression and noted the differences between the group that achieved virologic suppression early (<1 year of age; group 1) vs the group that achieved virologic suppression late (>1 year of age; group 2)[60]. An important observation was that on stratifying the data by age at virologic suppression, the decay rates for the proviral reservoir were different based on the age at virologic suppression. For the group 1 (VS<1 year), the rate of decay was much faster with vs group 2 (VS 1-5 years) with rates of -0.50 log_{10} HIV-1 DNA copies/million PBMCs and -0.15 log_{10} HIV-1 DNA copies/million PBMCs respectively. An inflexion point was observed at 2 years post confirmed VS, after which the slopes were not very different between the two groups with decay rates of -0.03 log_{10} HIV-1 DNA copies/million PBMCs and -0.05 log_{10} HIV-1 DNA copies/million PBMCs respectively[60]. This difference was noted even though both groups started off with similar concentrations of total HIV-1 DNA at the time of confirmed virologic suppression.

The observation of a faster decay rate of total HIV-1 DNA the group with earlier VS during the first 2 years of VS in the study by Uprety et al., sparked interest in understanding the dynamics of the reservoir and proviral species with timing and duration of VS[60]. However, as mentioned earlier, the study used a single-plex ddPCR assay to measure the total HIV-1 DNA load. Assays that measure

the total HIV-1 DNA overestimate the size of the reservoir as they also measure the defective proviruses which dominate the proviral landscape [64,65,68,70,144,145,151,153,155,165].

This overestimation of the reservoir size can negatively affect ART-free remission strategies aimed to clear the latent replication-competent reservoir, which is the reason for rebound of viremia when a person is taken off ART[48-53,87,88]. Hence in the current study, we applied the Intact Proviral DNA Assay (which can differentiate between intact and defective proviruses)[69], to study the composition of the reservoir and understand what led to a rapid decay of the reservoir size in early suppressed children[69].

For the current study on the PHACS cohort, we hypothesized that the group with early treatment and suppression ended up freezing the reservoir with a higher proportion of intact HIV-1 DNA (at the time of ART initiation) that would have potential for selective elimination compared to the cohort with later virologic suppression, where defective proviruses would preferentially accumulate and therefore be less likely to be cleared. The reason behind this hypothesis was late suppression, mimics untreated infection where the HIV-1 is able to circulate long enough for the development of HIV-1 specific immune response[57]. Hence when the intact proviruses are reactivated (late suppression), they are rapidly cleared by the immune response and viral cytopathic effects giving rise to a landscape dominated by defective proviruses. This would potentially explain why the decay rate for the total HIV-1 DNA would be faster within the first two years of suppression in group 1 vs group 2.

To test this hypothesis, samples from the PHACS-AMP cohort that were remnant from the previous study[60] were taken and tested with the IPDA to understand the composition and dynamics of the reservoir. These samples also helped in understanding the effect of duration of the early antiretroviral therapy and virologic suppression on the dynamics and composition of the reservoir over time.

In this study, it was noted that for some participants the total HIV-1 DNA was stable while the intact and defective HIV-1 DNA were in a flux. It was observed that a decline in intact HIV-1 DNA was compensated with increases in the proportions of 5' deleted or 3' deleted/hypermutated HIV-1 DNA. This was a very interesting observation as it could have implications in terms of reservoir maintenance.

What contributed or led to this proportional change in the defective HIV-1 genomes? Could clonal amplification have played a role in the change of the defective HIV-1 genomes? Clonal amplification has been previously speculated and shown to be involved in the maintenance of the defective genomes[57,69,111,167,176,232].

In the current study, 8% of the participants (2/25) demonstrated primer-probe mismatches, 4% showed a mismatch for the env region and 4% showed a mismatch for the gag (psi) region respectively. The next step would be to create new primer-probes for these participants and run the IPDA again on these samples to study the reservoir. In a study done by Simonetti et al, in 400 adults on effective ART, across several cohorts, 2.8% showed env signal failure and

primer-probe mismatch while 3.5% showed gag (psi) signal failure and primer-probe mismatch[171].

In another study by Gaebler et al., involving 39 participants on effective ART, 18% of the participants demonstrated IPDA gag (psi) or env signal failure[169]. Of these at least 2 belonged to another clade of HIV-1 group M while the remaining 14.7% belonged to clade B[169].

An interesting observation in the study by Gaebler et al., was that 4/39 participants i.e., 10.3% demonstrated reduced fluorescent signaling. The reason behind this was a mismatch in the env probe. Therefore, two different types of signal reductions were observed in adults with respect to mismatches in the primer-probe or just the probe[169]. In contrast, for the current study, only signal failure and not low fluorescence was observed. However, due to the small cohort size, it is possible that low fluorescence amplitude may not have caused a problem as yet but may do so when the IPDA is applied to a much larger cohort in pediatric participants[169].

One striking observation made in our cohort was seen in the case of participant PHACS sample 25. PHACS sample 25 demonstrated what appeared to be a double signal failure in the IPDA 2D amplitude. At first, we thought that this might be a case of primer-probe mismatch perhaps owing to subtype differences. The PCR results indicated otherwise, upon two tries, no amplicon bands for gag (psi) or env were retrieved. The PCR had worked for other samples, so it appeared to an issue with the sample. However, upon checking the POL-LTR assay results for the particular time point for PHACS sample 25, it was noted that the value for POL HIV-1 DNA copies/million PBMCs was also

undetectable. Owing to lack of sample quantity and only one time point for this participant, based on the IPDA and the POL-LTR assay, it was concluded that this participant may have had very low proviral loads that were undetectable by both of the assays[60]. Since only 200 ng of DNA was tested in 8 replicates, the number of cells analyzed fell between 150000 - 350000 which could have also led to this lack of detection of proviruses. Increasing the DNA input, and therefore the number of cells analyzed, may allow detection of the proviruses in this sample as suggested by a recent study done by Peluso et al[62].

For some participants, the intact HIV-1 DNA copies/million PBMCs did not stably decline over time, instead occasional fluctuations were observed. These samples need further analysis to understand the factors that lead to this state of flux. A viral blip may have occurred prior to sample collection which would explain the sudden increase in intact HIV- 1 DNA copies/million PBMCs in some cases. Another possibility could be high 2LTR circle values which are considered to be indicative of new infections, although this still remains a point of debate in the field[233-235].

A drawback of the IPDA, as mentioned earlier, is that it cannot differentiate between unintegrated and integrated forms of HIV-1 DNA and is not very accurate in measuring the reservoir during ongoing replication. Therefore, 2LTR circle values ≥ 10 copies/million PBMCs could have impacted the results on the individual basis and hence those time points were eliminated from further analysis.

Crude HLA-typing could also be done to determine if the sample indeed belonged to the respective participant. In such long-term studies, the samples

are collected at testing sites and undergo a series of processing steps before they arrive at the laboratories for analysis. In addition to this, depending on the number of tests, the sample could be handled several times and manual errors could result in mislabeling which could affect downstream analysis. Therefore, in situations where there is doubt with respect to HIV biology, crude HLA typing can be a beneficial tool to confirm that there was no sample mix up. Similar results were observed in the study by Peluso et al., where some participants demonstrated either no change in the intact HIV-1 DNA CPM PBMCs over time on ART or demonstrated an expansion in the intact HIV- 1 DNA CPM PBMCs[62].

Next, the data were stratified based on age at virologic suppression and duration of virologic suppression. Based on age at virologic suppression, as expected, the distribution of intact HIV-1 DNA proviruses was lower in the group that achieved virologic suppression (VS) at less than one year of age (Group 1) vs those that achieved virologic suppression between 1-5 years of age (Group 2). A few reasons could explain why group 2 had a higher distribution of the intact HIV-1 DNA copies had a higher distribution in group 2: Since we were blinded, we could not identify if the samples measured for group 1 were from the time points later in the virologic suppression (greater than 2 years of VS). In addition to this, group 2 samples could have been from earlier time points (less than 2 years of VS, 2-5years of VS). Also, there were more samples with detectable intact HIV-1 DNA copies (30/40), double the number of time points were available for this group vs group 1 with 40 and 21 time points respectively.

Another possibility for this result could be clonal expansion of the intact HIV-1 proviral reservoirs in the later suppressed group. Expansion of the intact HIV-1 genomes has been reported in adults although not at the same level as defective HIV-1 genomes[57,166]. A third possibility could be that early during the infection, the intact HIV-1 proviruses may have integrated into inactive genes which would prevent them from expressing viral proteins and therefore protect the virus from immune clearance. It could also be possible that with later suppression, the immune repertoire of the participants had a higher proportion of resting CD4 T cells than group 1, thereby allowing the virus to enter a more long-lived cell population[23,65].

The distribution of both 3' deleted/hypermutated and 5' deleted HIV-1 proviral genomes, was as expected with group 2 having a much higher concentration overall. The distribution of total HIV-1 DNA was also higher in the case of group 2 as would be expected for participants that suppressed at a later age. When the data were analyzed based on the first and last time points, the concentration of the defective (3' deleted/hypermutated and 5' deleted) HIV-1 proviruses in both the groups had increased between the first and the last time point, indicating that the defective genomes might have undergone clonal expansion. Integration site analysis needs to be performed to support this observation and quantify the expansion [57,69,111,167].

The intact HIV-1 DNA copies showed similar decay profiles between the two groups. In group 1, the intact HIV-1 DNA copies declined between the first and last time point, the median intact HIV-1 DNA copies/million PBMCs for both the time points were 4.9 and 4.6 copies/million PBMCs. The apparent slow decline

could be due to the poor detection of intact HIV-1 DNA copies, 1/21 at both time points. Another reason could be the longevity of the cells in which the intact HIV-1 proviral reservoir might be situated or the site of integration. In comparison to group 1, the intact HIV-1 DNA copies/million PBMCs for group 2 also declined between the first and last time point. It is possible that some of the intact HIV-1 proviruses in group 2 got reactivated and hence were cleared by the immune mechanisms or viral cytopathic effects.

Based on duration of suppression, the intact HIV-1 DNA copies/million PBMCs continued to decline from less than 2 years of VS to greater than 5 years of VS. The defective HIV-1 DNA copies/million followed similar patterns of decline followed by an increase as seen when the data were stratified by age at virologic control. The proportion of intact HIV-1 DNA that constitutes the total HIV- 1 DNA copies/million PBMCs declined from a median of 22.6% when the duration of VS<2 years to a median of 5.2% when the duration of suppression was greater than 5 years. The intact HIV-1 DNA demonstrated a decline between each group of duration of VS with medians of 41.2 (VS<2 years), 8.3 (VS=2-5 years) and 4.9 (VS>5 years) intact HIV-1 DNA copies/million PBMCs. The decline in intact HIV-1 DNA copies/million PMBCs between VS<2 years and VS 2- 5 years, and VS<2 years and VS>5 years were statistically significant. Although there were several undetectable values, our conservative approach towards assigning those values their respective limits of detection (based on number of cells analyzed), ensure that the change we noted was definitive.

This data further demonstrated that with increasing duration of suppression, the intact HIV-1 proviral reservoir is diminished, albeit slowly in some cases while the defective HIV-1 proviruses show expansion over time even in the case of early treated pediatric infections[57,189]. This study was able to show that the reservoir in pediatric cases is very dynamic. The size of the reservoir changes with age at VS as well as with duration of VS with the intact proviral reservoir being the most affected.

As the measure of differential decay of the reservoir from the previous studies was based on total HIV-1 DNA, it was important to understand the dynamics and proportion of the different species within the reservoir with respect to this change[59,60]. The current study helped bridge this gap by demonstrating that in most cases, it was the intact HIV-1 proviral reservoir that was primarily responsible for the decline observed in the previous study. This has implications in terms of the time at which remission strategies to target the reservoir can be applied to achieve ART-free remission.

The IPDA has been applied more broadly in cases of adult HIV-1 infections, where it was observed that intact proviruses decline with increasing duration of virologic suppression while defective proviruses are in a state of flux[57,62]. This is similar to what was observed in this study on pediatric HIV-1 infections.

However, differences arise when developing an ART-free remission strategy. In adult HIV-1 Infections, in most cases, ART is initiated during chronic infection[62,69,166,168,169,171] and even in the case of treatment during acute infections, the participants initiate ART at least 6 months after HIV-1

infection[56,58,61,168,183]. This allows the reservoir to establish itself, give rise to viral diversity as the virus circulates uncontrolled till ART is initiated[25].

In comparison, in perinatal HV-1 infection, treatment can be started within hours to days of birth[92] or 6 months post HIV-1 infection[55,59,60,191,200,201,203,205,215,216 222] which is unique to pediatrics. In non-human primates Simian Immunodeficiency Virus (SIV) is able to establish the reservoir within 3 days[236]. Although very early ART is unable to block reservoir establishment, it helps reduce the size of the reservoir by suppressing viremia early. The small size and low diversity of the reservoir in early treated perinatal HIV-1 infections make it suitable to develop and target strategies to achieve ART-free remission as compared to adults [59,63,198].

The current study was limited by the small cohort size, use of remnant genomic DNA which led to increased shearing, and lack of PBMC samples for similar time points across all participants. In spite of these limitations, insights were gained into the dynamics of the reservoir in perinatal HIV-1 Infections:

1. Early suppression was associated with lower distribution of intact HIV-1 proviruses implying that they may be cleared rapidly upon initiation of ART.
2. The intact HIV-1 proviral reservoir decreased with increased duration of effective ART although the defective HIV-1 proviruses may not have followed this trajectory.
3. Higher intact HIV-1 proviral copies in the later suppressed population could be due to clonal amplification or integration into inactive sites.
4. Integration site analysis will help clarify point 3 and the role clonal amplification may play in the maintenance of the defective HIV-1 genomes.

5. Perinatal HIV-1 reservoir dynamics follow similar trends to that seen in adult infections implying that early treatment may play a role in the trajectory of the reservoir size and composition, and the capacity to expand clonally.

The results of the IPDA should be validated with another assay such as near full-length genome sequencing or Quantitative Quadraplex PCR (Q4PCR) to ensure an even more accurate description of the reservoir size. [169]

Conclusion and Future Directions

Based on the results of this study, it is clear that early treatment and early suppression of HIV-1 replication in children with perinatal HIV-1 infections reduces the size of the intact proviral reservoir over time. However, further studies need to be done to answer the following questions:

At what rate does the intact proviral reservoir decay with early suppression vs late suppression?[57,60,62,166] How much of the intact proviral reservoir detected by IPDA can be induced in-vivo?[151] This is important as intact proviruses that have been latently silenced or are not replication competent will continue to appear as intact in the IPDA, thereby affecting the results of potential ART-free remission strategies. What threshold of reduction in the size of the intact proviral reservoir can be used in perinatal infections to determine the effectiveness of an ART-free remission strategy?[168] In a study on adults living with HIV-1, on ART, it was found that a greater than 6-fold decrease in the size of the replication competent proviruses was a reliable metric to determine the effects of the ART-free remission strategies[90]. Will this threshold also apply to the reservoir children?

As mentioned earlier, pediatric and adult infections may follow the same trajectory but are different in terms of: 1) the size of the reservoir; 2) the % intact detectable after long-term suppression which is 1% in pediatric cases vs 2% in adults[195]; 3) and the immune repertoire which is reconstituted by naive CD4 T cells in children[30] which dilute the replication competent reservoir. Therefore, studies on the reservoir need to continue in perinatal HIV-1 infections.

This preliminary study had several limitations, therefore a larger study with more time points, fresh genomic DNA samples and similar time points would be beneficial in answering some of these questions.

The data from the larger study could be analyzed to understand the effect of age at virologic suppression on the duration of the virologic suppression and the subsequently the size of the reservoir. Additional experiments such as developing specific primer probes for participants that initially demonstrated primer-probe mismatches could be done to probe the reservoir in those participants.

Future studies in perinatal HIV-1 infections could include applying the IPDA across several cohorts to understand the performance characteristics of the assay similar to the study by Simonetti et al[171]. Near full-length genome sequencing (nFGS) or matched integration site proviral genome sequencing (MIP-seq) could be performed to elucidate 1) the role of clonal expansion in maintenance of the reservoir in perinatal HIV-1 infections and 2) the preferential sites of integration for intact proviruses in perinatal HIV-1 infections to understand the potential for reactivation of replication-competent proviruses[174,177].

Studies on the DNA methylation pattern could also be done to understand whether a provirus classified as intact is capable of reactivation or has been silenced (block)[237,238].

In summary, the current preliminary study showed that intact proviruses are selectively eliminated in the early suppressed children vs late suppressed children which explains why the decay rates of the total HIV-1 DNA were different between the two groups post ART initiation. This could help elucidate an appropriate time to administer remission strategies in addition to ART to help eradicate the replication-competent reservoir. However, more research is needed to truly understand the decay dynamics of the reservoir in perinatal HIV-1 infections to begin working towards a functional cure/remission.

References

1. Gallo RC, Wong-Staal F. A Human T-Lymphotropic Retrovirus (HTLV-III) as the Cause of the Acquired Immunodeficiency Syndrome. *Journal of Urology.* 1986;136(2):548-548.
2. Gallo RC, Montagnier, L. The Discovery of HIV as the Cause of AIDS. *N Engl J Med* 2003;349(24).
3. Luzuriaga K, Mofenson LM. Challenges in the Elimination of Pediatric HIV-1 Infection. *N Engl J Med.* 2016;374(8):761-770.
4. Popovic M, M.G. S, Read E, Gallo RC. Detection, Isolation, and Continuous Production of Cytopathic Retroviruses (HTLV-111) from Patients with AIDS and Pre-AIDS. *Science.* 1984;224.
5. F. Barre-Sinoussi JCCFR, M. T. Nugeyre S. Chamaret, J. Gruest C. Dauguet, C. Axler-Blin, F. Vtzinet-Brun, C. Rouzioux,W. Rozenbaum,L. Montagnier. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science.* 1983;220.
6. Fields Bernard N DMK, Peter M.Howley and Diane E Griffin. . *Field Virology (Specific Virus Families Retroviridae).* 6th ed. ed: Philadelphia Lippincott Williams & Wilkin.; 2013.
7. Luzuriaga K. Early Combination Antiretroviral Therapy Limits HIV-1 Persistence in Children. *Annu Rev Med.* 2016;67:201-213.
8. Wilfert CM, Wilson C, Luzuriaga K, Epstein L. Pathogenesis of Pediatric Human-Immunodeficiency-Virus Type-1 Infection. *Journal of Infectious Diseases.* 1994;170(2):286-292.
9. Broder S, Gallo RC. A pathogenic retrovirus (HTLV-III) linked to AIDS. *N Engl J Med.* 1984;311(20):1292-1297.
10. Bryson YJ, Luzuriaga K, Sullivan JL, Wara DW. Proposed definitions for in utero versus intrapartum transmission of HIV-1. *N Engl J Med.* 1992;327(17):1246-1247.
11. Wara DW, Luzuriaga K, Martin NL, Sullivan JL, Bryson YJ. Maternal transmission and diagnosis of human immunodeficiency virus during infancy. *Ann N Y Acad Sci.* 1993;693:14-19.
12. Davey RT, Jr., Bhat N, Yoder C, et al. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proc Natl Acad Sci U S A.* 1999;96(26):15109-15114.
13. Vella S, Schwartlander B, Sow SP, Eholie SP, Murphy RL. The history of antiretroviral therapy and of its implementation in resource-limited areas of the world. *AIDS.* 2012;26(10):1231-1241.
14. Wei X, Ghosh SK, Taylor ME, et al. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature.* 1995;373(6510):117-122.
15. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature.* 1995;373(6510):123-126.
16. Perelson AS, Essunger P, Cao Y, et al. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature.* 1997;387(6629):188-191.
17. Hammer SM, Squires KE, Hughes MD, et al. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and

CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med.* 1997;337(11):725-733.
18. Gulick RM, Mellors JW, Havlir D, et al. Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med.* 1997;337(11):734-739.
19. Piatak M, Jr., Saag MS, Yang LC, et al. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science.* 1993;259(5102):1749-1754.
20. UNAIDS. Global HIV & AIDS statistics — 2020 fact sheet https://www.unaids.org/en/resources/fact-sheet. Published 2020. Accessed 12/09/2020.
21. Luzuriaga K, Sullivan JL. Prevention and treatment of pediatric HIV infection. *Jama-J Am Med Assoc.* 1998;280(1):17-18.
22. Blankson JN, Persaud D, Siliciano RF. The challenge of viral reservoirs in HIV-1 infection. *Annual Review of Medicine.* 2002;53:557-593.
23. Eisele E, Siliciano RF. Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity.* 2012;37(3):377-388.
24. Ho YC, Siliciano JD. Efforts to eliminate the latent reservoir in resting CD4+ T cells: strategies for curing HIV-1 infection. *J Virus Erad.* 2015;1(4):229-231.
25. Ferguson MR, Rojo DR, von Lindern JJ, O'Brien WA. HIV-1 replication cycle. *Clin Lab Med.* 2002;22(3):611-635.
26. Jane Flint VRR, Glenn F. Rall, Anna Marie Skalka, and Lynn W. Enquist. *Principles of Virology.* Vol II. 4th Ed. ed: ASM Press, Washington D.C.; 2015.
27. Li C, Burdick RC, Nagashima K, Hu WS, Pathak VK. HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proc Natl Acad Sci U S A.* 2021;118(10).
28. Jacques DA, McEwan WA, Hilditch L, Price AJ, Towers GJ, James LC. HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature.* 2016;536(7616):349-353.
29. Temin HM. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc Natl Acad Sci U S A.* 1993;90(15):6900-6903.
30. Goulder PJ, Lewin SR, Leitman EM. Paediatric HIV infection: the potential for cure. *Nat Rev Immunol.* 2016;16(4):259-271.
31. Mold JE, Michaelsson J, Burt TD, et al. Maternal alloantigens promote the development of tolerogenic fetal regulatory T cells in utero. *Science.* 2008;322(5907):1562-1565.
32. Levy O, Coughlin M, Cronstein BN, Roy RM, Desai A, Wessels MR. The adenosine system selectively inhibits TLR-mediated TNF-alpha production in the human newborn. *J Immunol.* 2006;177(3):1956-1966.
33. Upham JW, Lee PT, Holt BJ, et al. Development of interleukin-12-producing capacity throughout childhood. *Infect Immun.* 2002;70(12):6583-6588.
34. Kollmann TR, Crabtree J, Rein-Weston A, et al. Neonatal innate TLR-mediated responses are distinct from those of adults. *J Immunol.* 2009;183(11):7150-7160.
35. Elyaman W, Bradshaw EM, Uyttenhove C, et al. IL-9 induces differentiation of TH17 cells and enhances function of FoxP3+ natural regulatory T cells. *Proc Natl Acad Sci U S A.* 2009;106(31):12885-12890.
36. Bunders MJ, van der Loos CM, Klarenbeek PL, et al. Memory CD4(+)CCR5(+) T cells are abundantly present in the gut of newborn infants to facilitate mother-to-child transmission of HIV-1. *Blood.* 2012;120(22):4383-4390.
37. Gosselin A, Monteiro P, Chomont N, et al. Peripheral blood CCR4+CCR6+ and CXCR3+CCR6+CD4+ T cells are highly permissive to HIV-1 infection. *J Immunol.* 2010;184(3):1604-1616.

38. Monteiro P, Gosselin A, Wacleche VS, et al. Memory CCR6+CD4+ T cells are preferential targets for productive HIV type 1 infection regardless of their expression of integrin beta7. *J Immunol.* 2011;186(8):4618-4630.
39. Fischl MA, Richman DD, Grieco MH, et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *N Engl J Med.* 1987;317(4):185-191.
40. Johnson VA, Merrill DP, Videler JA, et al. Two-drug combinations of zidovudine, didanosine, and recombinant interferon-alpha A inhibit replication of zidovudine-resistant human immunodeficiency virus type 1 synergistically in vitro. *J Infect Dis.* 1991;164(4):646-655.
41. Eron JJ, Jr., Johnson VA, Merrill DP, Chou TC, Hirsch MS. Synergistic inhibition of replication of human immunodeficiency virus type 1, including that of a zidovudine-resistant isolate, by zidovudine and 2',3'-dideoxycytidine in vitro. *Antimicrob Agents Chemother.* 1992;36(7):1559-1562.
42. Cheeseman SH, Havlir D, McLaughlin MM, et al. Phase I/II evaluation of nevirapine alone and in combination with zidovudine for infection with human immunodeficiency virus. *J Acquir Immune Defic Syndr Hum Retrovirol.* 1995;8(2):141-151.
43. Schooley RT, Campbell TB, Kuritzkes DR, et al. Phase 1 study of combination therapy with L-697,661 and zidovudine. The ACTG 184 Protocol Team. *J Acquir Immune Defic Syndr Hum Retrovirol.* 1996;12(4):363-370.
44. Tremblay C, Merrill DP, Chou TC, Hirsch MS. Interactions among combinations of two and three protease inhibitors against drug-susceptible and drug-resistant HIV-1 isolates. *J Acquir Immune Defic Syndr.* 1999;22(5):430-436.
45. Campbell TB, Young RK, Eron JJ, D'Aquila RT, Tarpley WG, Kuritzkes DR. Inhibition of human immunodeficiency virus type 1 replication in vitro by the bisheteroarylpiperazine atevirdine (U-87201E) in combination with zidovudine or didanosine. *J Infect Dis.* 1993;168(2):318-326.
46. Collier AC, Coombs RW, Fischl MA, et al. Combination therapy with zidovudine and didanosine compared with zidovudine alone in HIV-1 infection. *Ann Intern Med.* 1993;119(8):786-793.
47. Caliendo AM, Hirsch MS. Combination therapy for infection due to human immunodeficiency virus type 1. *Clin Infect Dis.* 1994;18(4):516-524.
48. Wong JK, Hezareh M, Gunthard HF, et al. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science.* 1997;278(5341):1291-1295.
49. Chun TW, Stuyver L, Mizell SB, et al. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *P Natl Acad Sci USA.* 1997;94(24):13193-13197.
50. Siliciano JD, Kajdas J, Finzi D, et al. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med.* 2003;9(6):727-728.
51. Finzi D, Blankson J, Siliciano JD, et al. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med.* 1999;5(5):512-517.
52. Finzi D, Hermankova M, Pierson T, et al. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science.* 1997;278(5341):1295-1300.
53. Chun TW, Davey RT, Jr., Engel D, Lane HC, Fauci AS. Re-emergence of HIV after stopping therapy. *Nature.* 1999;401(6756):874-875.
54. Ananworanich J, Dube K, Chomont N. How does the timing of antiretroviral therapy initiation in acute infection affect HIV reservoirs? *Curr Opin HIV AIDS.* 2015;10(1):18-28.

55. Persaud D, Palumbo PE, Ziemniak C, et al. Dynamics of the resting CD4(+) T-cell latent HIV reservoir in infants initiating HAART less than 6 months of age. *AIDS.* 2012;26(12):1483-1490.
56. Jain V, Hartogensis W, Bacchetti P, et al. Antiretroviral therapy initiated within 6 months of HIV infection is associated with lower T-cell activation and smaller HIV reservoir size. *J Infect Dis.* 2013;208(8):1202-1211.
57. Gandhi RT, Cyktor JC, Bosch RJ, et al. Selective Decay of Intact HIV-1 Proviral DNA on Antiretroviral Therapy. *J Infect Dis.* 2021;223(2):225-233.
58. Hocqueloux L, Avettand-Fenoel V, Jacquot S, et al. Long-term antiretroviral therapy initiated during primary HIV-1 infection is key to achieving both low HIV reservoirs and normal T cell counts. *J Antimicrob Chemother.* 2013;68(5):1169-1178.
59. Persaud D, Patel K, Karalius B, et al. Influence of age at virologic control on peripheral blood human immunodeficiency virus reservoir size and serostatus in perinatally infected adolescents. *JAMA Pediatr.* 2014;168(12):1138-1146.
60. Uprety P, Patel K, Karalius B, et al. Human Immunodeficiency Virus Type 1 DNA Decay Dynamics With Early, Long-term Virologic Control of Perinatal Infection. *Clin Infect Dis.* 2017;64(11):1471-1478.
61. Chun TW, Justement JS, Moir S, et al. Decay of the HIV reservoir in patients receiving antiretroviral therapy for extended periods: implications for eradication of virus. *J Infect Dis.* 2007;195(12):1762-1764.
62. Peluso MJ, Bacchetti P, Ritter KD, et al. Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy. *JCI Insight.* 2020;5(4).
63. Persaud D, Ray SC, Kajdas J, et al. Slow human immunodeficiency virus type 1 evolution in viral reservoirs in infants treated with effective antiretroviral therapy. *AIDS Res Hum Retroviruses.* 2007;23(3):381-390.
64. Abdel-Mohsen M, Richman D, Siliciano RF, et al. Recommendations for measuring HIV reservoir size in cure-directed clinical trials. *Nat Med.* 2020;26(9):1339-1350.
65. Barton KM, Palmer SE. How to Define the Latent Reservoir: Tools of the Trade. *Curr HIV/AIDS Rep.* 2016;13(2):77-84.
66. Pace MJ, Agosto L, Graf EH, O'Doherty U. HIV reservoirs and latency models. *Virology.* 2011;411(2):344-354.
67. Siliciano RF, Greene WC. HIV latency. *Cold Spring Harb Perspect Med.* 2011;1(1):a007096.
68. Siliciano JD, Siliciano RF. Assays to Measure Latency, Reservoirs, and Reactivation. *Curr Top Microbiol Immunol.* 2018;417:23-41.
69. Bruner KM, Wang Z, Simonetti FR, et al. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature.* 2019;566(7742):120-125.
70. Bruner KM, Murray AJ, Pollack RA, et al. Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat Med.* 2016;22(9):1043-1049.
71. Imamichi H, Dewar RL, Adelsberger JW, et al. Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. *Proc Natl Acad Sci U S A.* 2016;113(31):8783-8788.
72. Pollack RA, Jones RB, Pertea M, et al. Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. *Cell Host Microbe.* 2017;21(4):494-506 e494.
73. Ho YC, Shan L, Hosmane NN, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell.* 2013;155(3):540-551.
74. Persaud D, Pierson T, Ruff C, et al. A stable latent reservoir for HIV-1 in resting CD4(+) T lymphocytes in infected children. *J Clin Invest.* 2000;105(7):995-1003.

75. Dieffenbach CW, Fauci AS. Thirty years of HIV and AIDS: future challenges and opportunities. *Ann Intern Med.* 2011;154(11):766-771.
76. Rainwater-Lovett K, Luzuriaga K, Persaud D. Very early combination antiretroviral therapy in infants: prospects for cure. *Curr Opin HIV AIDS.* 2015;10(1):4-11.
77. Chun TW, Carruth L, Finzi D, et al. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature.* 1997;387(6629):183-188.
78. Chomont N, El-Far M, Ancuta P, et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med.* 2009;15(8):893-900.
79. Coiras M, Lopez-Huertas MR, Perez-Olmeda M, Alcami J. Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat Rev Microbiol.* 2009;7(11):798-812.
80. Dahl V, Palmer S. Establishment of drug-resistant HIV-1 in latent reservoirs. *J Infect Dis.* 2009;199(9):1258-1260.
81. Honeycutt JB, Wahl A, Baker C, et al. Macrophages sustain HIV replication in vivo independently of T cells. *J Clin Invest.* 2016;126(4):1353-1366.
82. Igarashi T, Brown CR, Endo Y, et al. Macrophage are the principal reservoir and sustain high virus loads in rhesus macaques after the depletion of CD4+ T cells by a highly pathogenic simian immunodeficiency virus/HIV type 1 chimera (SHIV): Implications for HIV-1 infections of humans. *Proc Natl Acad Sci U S A.* 2001;98(2):658-663.
83. Calantone N, Wu F, Klase Z, et al. Tissue myeloid cells in SIV-infected primates acquire viral DNA through phagocytosis of infected T cells. *Immunity.* 2014;41(3):493-502.
84. Gama L, Abreu CM, Shirk EN, et al. Reactivation of simian immunodeficiency virus reservoirs in the brain of virally suppressed macaques. *AIDS.* 2017;31(1):5-14.
85. Chahroudi A, Wagner TA, Persaud D. CNS Persistence of HIV-1 in Children: the Untapped Reservoir. *Curr HIV/AIDS Rep.* 2018;15(5):382-387.
86. Wang Z, Simonetti FR, Siliciano RF, Laird GM. Measuring replication competent HIV-1: advances and challenges in defining the latent reservoir. *Retrovirology.* 2018;15(1):21.
87. Davey RT, Bhat N, Yoder C, et al. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *P Natl Acad Sci USA.* 1999;96(26):15109-15114.
88. Chun TW, Finzi D, Margolick J, Chadwick K, Schwartz D, Siliciano RF. In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nat Med.* 1995;1(12):1284-1290.
89. Hill AL, Rosenbloom DI, Fu F, Nowak MA, Siliciano RF. Predicting the outcomes of treatment to eradicate the latent reservoir for HIV-1. *Proc Natl Acad Sci U S A.* 2014;111(37):13475-13480.
90. Crooks AM, Bateson R, Cope AB, et al. Precise Quantitation of the Latent HIV-1 Reservoir: Implications for Eradication Strategies. *J Infect Dis.* 2015;212(9):1361-1365.
91. Strain MC, Gunthard HF, Havlir DV, et al. Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic stability predicts lifelong persistence. *Proc Natl Acad Sci U S A.* 2003;100(8):4819-4824.
92. Persaud D, Gay H, Ziemniak C, et al. Absence of detectable HIV-1 viremia after treatment cessation in an infant. *N Engl J Med.* 2013;369(19):1828-1835.
93. Kaufmann GR, Khanna N, Weber R, et al. Long-term virological response to multiple sequential regimens of highly active antiretroviral therapy for HIV infection. *Antivir Ther.* 2004;9(2):263-274.
94. Luzuriaga K, Gay H, Ziemniak C, et al. Viremic relapse after HIV-1 remission in a perinatally infected child. *N Engl J Med.* 2015;372(8):786-788.
95. Larder BA, Darby G, Richman DD. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science.* 1989;243(4899):1731-1734.

96. Sengupta S, Siliciano RF. Targeting the Latent Reservoir for HIV-1. *Immunity.* 2018;48(5):872-895.
97. UNAIDS. *UNAIDS Fact Sheet: GLOBAL HIV STATISTICS.* 2019.
98. Cohn LB, Chomont N, Deeks SG. The Biology of the HIV-1 Latent Reservoir and Implications for Cure Strategies. *Cell Host Microbe.* 2020;27(4):519-530.
99. Kenneth Murphy CW. *Janeway's Immunobiology (Integrated Dynamics of Innate and Adaptive Immunity).* 9th ed. New York: Garland Science, Taylor & Francis Group, LLC; 2017.
100. Michie CA, McLean A, Alcock C, Beverley PC. Lifespan of human lymphocyte subsets defined by CD45 isoforms. *Nature.* 1992;360(6401):264-265.
101. McLean AR, Michie CA. In vivo estimates of division and death rates of human T lymphocytes. *Proc Natl Acad Sci U S A.* 1995;92(9):3707-3711.
102. Hellerstein MK, Hoh RA, Hanley MB, et al. Subpopulations of long-lived and short-lived T cells in advanced HIV-1 infection. *J Clin Invest.* 2003;112(6):956-966.
103. Rausch JW, Le Grice SFJ. Characterizing the Latent HIV-1 Reservoir in Patients with Viremia Suppressed on cART: Progress, Challenges, and Opportunities. *Curr HIV Res.* 2020;18(2):99-113.
104. Katlama C, Lambert-Niclot S, Assoumou L, et al. Treatment intensification followed by interleukin-7 reactivates HIV without reducing total HIV DNA: a randomized trial. *AIDS.* 2016;30(2):221-230.
105. Vandergeeten C, Fromentin R, DaFonseca S, et al. Interleukin-7 promotes HIV persistence during antiretroviral therapy. *Blood.* 2013;121(21):4321-4329.
106. Purton JF, Tan JT, Rubinstein MP, Kim DM, Sprent J, Surh CD. Antiviral CD4+ memory T cells are IL-15 dependent. *J Exp Med.* 2007;204(4):951-961.
107. Bosque A, Famiglietti M, Weyrich AS, Goulston C, Planelles V. Homeostatic proliferation fails to efficiently reactivate HIV-1 latently infected central memory CD4+ T cells. *PLoS Pathog.* 2011;7(10):e1002288.
108. Henrich TJ, Hobbs KS, Hanhauser E, et al. Human Immunodeficiency Virus Type 1 Persistence Following Systemic Chemotherapy for Malignancy. *J Infect Dis.* 2017;216(2):254-262.
109. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell.* 2002;110(4):521-529.
110. Han Y, Lassen K, Monie D, et al. Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *J Virol.* 2004;78(12):6122-6133.
111. Maldarelli F, Wu X, Su L, et al. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science.* 2014;345(6193):179-183.
112. Wagner TA, McLaughlin S, Garg K, et al. HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science.* 2014;345(6196):570-573.
113. Cohn LB, Silva IT, Oliveira TY, et al. HIV-1 integration landscape during latent and active infection. *Cell.* 2015;160(3):420-432.
114. Lorenzi JC, Cohen YZ, Cohn LB, et al. Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA. *Proc Natl Acad Sci U S A.* 2016;113(49):E7908-E7916.
115. Bui JK, Sobolewski MD, Keele BF, et al. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS Pathog.* 2017;13(3):e1006283.
116. Hosmane NN, Kwon KJ, Bruner KM, et al. Proliferation of latently infected CD4(+) T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J Exp Med.* 2017;214(4):959-972.

117. Simonetti FR, Sobolewski MD, Fyne E, et al. Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo. *Proc Natl Acad Sci U S A.* 2016;113(7):1883-1888.
118. O'Doherty U. Mechanisms of human immunodeficiency virus-1 latency. *Transfusion.* 2005;45(2 Suppl):88S-91S.
119. Coull JJ, Romerio F, Sun JM, et al. The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. *J Virol.* 2000;74(15):6790-6799.
120. Hsia SC, Shi YB. Chromatin disruption and histone acetylation in regulation of the human immunodeficiency virus type 1 long terminal repeat by thyroid hormone receptor. *Mol Cell Biol.* 2002;22(12):4043-4052.
121. Chan JK, Greene WC. NF-kappaB/Rel: agonist and antagonist roles in HIV-1 latency. *Curr Opin HIV AIDS.* 2011;6(1):12-18.
122. Williams SA, Chen LF, Kwon H, Ruiz-Jarabo CM, Verdin E, Greene WC. NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *EMBO J.* 2006;25(1):139-149.
123. Kauder SE, Bosque A, Lindqvist A, Planelles V, Verdin E. Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog.* 2009;5(6):e1000495.
124. Karn J. The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit. *Curr Opin HIV AIDS.* 2011;6(1):4-11.
125. Ne E, Palstra RJ, Mahmoudi T. Transcription: Insights From the HIV-1 Promoter. *Int Rev Cell Mol Biol.* 2018;335:191-243.
126. Nabel G, Baltimore D. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature.* 1987;326(6114):711-713.
127. Kinoshita S, Chen BK, Kaneshima H, Nolan GP. Host control of HIV-1 parasitism in T cells by the nuclear factor of activated T cells. *Cell.* 1998;95(5):595-604.
128. Verdin E, Paras P, Jr., Van Lint C. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *EMBO J.* 1993;12(8):3249-3259.
129. Cary DC, Fujinaga K, Peterlin BM. Molecular mechanisms of HIV latency. *J Clin Invest.* 2016;126(2):448-454.
130. Shan L, Yang HC, Rabi SA, et al. Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. *J Virol.* 2011;85(11):5384-5393.
131. Lenasi T, Contreras X, Peterlin BM. Transcriptional interference antagonizes proviral gene expression to promote HIV latency. *Cell Host Microbe.* 2008;4(2):123-133.
132. Han Y, Lin YB, An W, et al. Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe.* 2008;4(2):134-146.
133. Lehrman G, Hogue IB, Palmer S, et al. Depletion of latent HIV-1 infection in vivo: a proof-of-concept study. *Lancet.* 2005;366(9485):549-555.
134. Prins JM, Jurriaans S, van Praag RM, et al. Immuno-activation with anti-CD3 and recombinant human IL-2 in HIV-1-infected patients on potent antiretroviral therapy. *AIDS.* 1999;13(17):2405-2410.
135. Kulkosky J, Nunnari G, Otero M, et al. Intensification and stimulation therapy for human immunodeficiency virus type 1 reservoirs in infected persons receiving virally suppressive highly active antiretroviral therapy. *J Infect Dis.* 2002;186(10):1403-1411.
136. Bouchat S, Delacourt N, Kula A, et al. Sequential treatment with 5-aza-2'-deoxycytidine and deacetylase inhibitors reactivates HIV-1. *EMBO Mol Med.* 2016;8(2):117-138.
137. Archin NM, Liberty AL, Kashuba AD, et al. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature.* 2012;487(7408):482-485.

138. Pache L, Dutra MS, Spivak AM, et al. BIRC2/cIAP1 Is a Negative Regulator of HIV-1 Transcription and Can Be Targeted by Smac Mimetics to Promote Reversal of Viral Latency. *Cell Host Microbe.* 2015;18(3):345-353.
139. Besnard E, Hakre S, Kampmann M, et al. The mTOR Complex Controls HIV Latency. *Cell Host Microbe.* 2016;20(6):785-797.
140. Mousseau G, Kessing CF, Fromentin R, Trautmann L, Chomont N, Valente ST. The Tat Inhibitor Didehydro-Cortistatin A Prevents HIV-1 Reactivation from Latency. *mBio.* 2015;6(4):e00465.
141. Kessing CF, Nixon CC, Li C, et al. In Vivo Suppression of HIV Rebound by Didehydro-Cortistatin A, a "Block-and-Lock" Strategy for HIV-1 Treatment. *Cell Rep.* 2017;21(3):600-611.
142. Horwitz JA, Bar-On Y, Lu CL, et al. Non-neutralizing Antibodies Alter the Course of HIV-1 Infection In Vivo. *Cell.* 2017;170(4):637-648 e610.
143. Kreitman RJ. Immunotoxins for targeted cancer therapy. *AAPS J.* 2006;8(3):E532-551.
144. Bruner KM, Hosmane NN, Siliciano RF. Towards an HIV-1 cure: measuring the latent reservoir. *Trends Microbiol.* 2015;23(4):192-203.
145. Strain MC, Lada SM, Luong T, et al. Highly precise measurement of HIV DNA by droplet digital PCR. *PLoS One.* 2013;8(4):e55943.
146. Ho DD, Moudgil T, Alam M. Quantitation of human immunodeficiency virus type 1 in the blood of infected persons. *N Engl J Med.* 1989;321(24):1621-1625.
147. Beliakova-Bethell N, Hezareh M, Wong JK, et al. Relative efficacy of T cell stimuli as inducers of productive HIV-1 replication in latently infected CD4 lymphocytes from patients on suppressive cART. *Virology.* 2017;508:127-133.
148. Siliciano JD, Siliciano RF. Enhanced culture assay for detection and quantitation of latently infected, resting CD4+ T-cells carrying replication-competent virus in HIV-1-infected individuals. *Methods Mol Biol.* 2005;304:3-15.
149. Archin NM, Kirchherr JL, Sung JA, et al. Interval dosing with the HDAC inhibitor vorinostat effectively reverses HIV latency. *J Clin Invest.* 2017;127(8):3126-3135.
150. Laird GM, Rosenbloom DI, Lai J, Siliciano RF, Siliciano JD. Measuring the Frequency of Latent HIV-1 in Resting CD4(+) T Cells Using a Limiting Dilution Coculture Assay. *Methods Mol Biol.* 2016;1354:239-253.
151. Falcinelli SD, Ceriani C, Margolis DM, Archin NM. New Frontiers in Measuring and Characterizing the HIV Reservoir. *Front Microbiol.* 2019;10:2878.
152. Laird GM, Eisele EE, Rabi SA, et al. Rapid quantification of the latent reservoir for HIV-1 using a viral outgrowth assay. *PLoS Pathog.* 2013;9(5):e1003398.
153. Eriksson S, Graf EH, Dahl V, et al. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathog.* 2013;9(2):e1003174.
154. Shan L, Rabi SA, Laird GM, et al. A novel PCR assay for quantification of HIV-1 RNA. *J Virol.* 2013;87(11):6521-6525.
155. Massanella M, Richman DD. Measuring the latent reservoir in vivo. *J Clin Invest.* 2016;126(2):464-472.
156. Imamichi H, Smith M, Adelsberger JW, et al. Defective HIV-1 proviruses produce viral proteins. *Proc Natl Acad Sci U S A.* 2020;117(7):3704-3710.
157. Rouzioux C, Melard A, Avettand-Fenoel V. Quantification of total HIV1-DNA in peripheral blood mononuclear cells. *Methods Mol Biol.* 2014;1087:261-270.
158. O'Doherty U, Swiggard WJ, Jeyakumar D, McGain D, Malim MH. A sensitive, quantitative assay for human immunodeficiency virus type 1 integration. *J Virol.* 2002;76(21):10942-10950.
159. Yu JJ, Wu TL, Liszewski MK, et al. A more precise HIV integration assay designed to detect small differences finds lower levels of integrated DNA in HAART treated patients. *Virology.* 2008;379(1):78-86.

160. Liszewski MK, Yu JJ, O'Doherty U. Detecting HIV-1 integration by repetitive-sampling Alu-gag PCR. *Methods.* 2009;47(4):254-260.
161. Mexas AM, Graf EH, Pace MJ, et al. Concurrent measures of total and integrated HIV DNA monitor reservoirs and ongoing replication in eradication trials. *AIDS.* 2012;26(18):2295-2306.
162. Sanchez G, Xu X, Chermann JC, Hirsch I. Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. *J Virol.* 1997;71(3):2233-2240.
163. Brady T, Kelly BJ, Male F, et al. Quantitation of HIV DNA integration: effects of differential integration site distributions on Alu-PCR assays. *J Virol Methods.* 2013;189(1):53-57.
164. Vandergeeten C, Fromentin R, Merlini E, et al. Cross-clade ultrasensitive PCR-based assays to measure HIV persistence in large-cohort studies. *J Virol.* 2014;88(21):12385-12396.
165. Lambrechts L, Cole B, Rutsaert S, Trypsteen W, Vandekerckhove L. Emerging PCR-Based Techniques to Study HIV-1 Reservoir Persistence. *Viruses.* 2020;12(2).
166. Antar AAR, Jenike KM, Jang S, et al. Longitudinal study reveals HIV-1-infected CD4+ T cell dynamics during long-term antiretroviral therapy. *J Clin Invest.* 2020.
167. Anderson EM, Maldarelli F. The role of integration and clonal expansion in HIV infection: live long and prosper. *Retrovirology.* 2018;15(1):71.
168. Falcinelli SD, Kilpatrick KW, Read J, et al. Longitudinal dynamics of intact HIV proviral DNA and outgrowth virus frequencies in a cohort of ART-treated individuals. *J Infect Dis.* 2020.
169. Gaebler C, Falcinelli SD, Stoffel E, et al. Sequence Evaluation and Comparative Analysis of Novel Assays for Intact Proviral HIV-1 DNA. *J Virol.* 2021;95(6).
170. Bender AM, Simonetti FR, Kumar MR, et al. The Landscape of Persistent Viral Genomes in ART-Treated SIV, SHIV, and HIV-2 Infections. *Cell Host Microbe.* 2019;26(1):73-85 e74.
171. Simonetti FR, White JA, Tumiotto C, et al. Intact proviral DNA assay analysis of large cohorts of people with HIV provides a benchmark for the frequency and composition of persistent proviral DNA. *Proc Natl Acad Sci U S A.* 2020;117(31):18692-18700.
172. Kinloch NN, Ren Y, Alberto WDC, et al. HIV Diversity Considerations in the Application of the Intact Proviral DNA Assay (IPDA). *bioRxiv.* 2020:2020.2005.2026.115006.
173. Gaebler C, Lorenzi JCC, Oliveira TY, et al. Combination of quadruplex qPCR and next-generation sequencing for qualitative and quantitative analysis of the HIV-1 latent reservoir. *J Exp Med.* 2019;216(10):2253-2264.
174. Patro SC, Brandt LD, Bale MJ, et al. Combined HIV-1 sequence and integration site analysis informs viral dynamics and allows reconstruction of replicating viral ancestors. *Proc Natl Acad Sci U S A.* 2019;116(51):25891-25899.
175. Thomas J, Ruggiero A, Paxton WA, Pollakis G. Measuring the Success of HIV-1 Cure Strategies. *Front Cell Infect Microbiol.* 2020;10:134.
176. Lee GQ, Orlova-Fink N, Einkauf K, et al. Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. *J Clin Invest.* 2017;127(7):2689-2696.
177. Einkauf KB, Lee GQ, Gao C, et al. Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy. *J Clin Invest.* 2019;129(3):988-998.
178. Procopio FA, Fromentin R, Kulpa DA, et al. A Novel Assay to Measure the Magnitude of the Inducible Viral Reservoir in HIV-infected Individuals. *EBioMedicine.* 2015;2(8):874-883.
179. Luzuriaga K, Persaud D. Treatment interruption after early-treated perinatal HIV-1 infection. *AIDS.* 2016;30(15):2381-2383.

180. Lassen KG, Ramyar KX, Bailey JR, Zhou Y, Siliciano RF. Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. *PLoS Pathog.* 2006;2(7):e68.
181. Lin X, Irwin D, Kanazawa S, et al. Transcriptional profiles of latent human immunodeficiency virus in infected individuals: effects of Tat on the host and reservoir. *J Virol.* 2003;77(15):8227-8236.
182. Razooky BS, Pai A, Aull K, Rouzine IM, Weinberger LS. A hardwired HIV latency program. *Cell.* 2015;160(5):990-1001.
183. Archin NM, Vaidya NK, Kuruc JD, et al. Immediate antiviral therapy appears to restrict resting CD4+ cell HIV-1 infection without accelerating the decay of latent infection. *Proc Natl Acad Sci U S A.* 2012;109(24):9523-9528.
184. Buzon MJ, Martin-Gayo E, Pereyra F, et al. Long-term antiretroviral treatment initiated at primary HIV-1 infection affects the size, composition, and decay kinetics of the reservoir of HIV-1-infected CD4 T cells. *J Virol.* 2014;88(17):10056-10065.
185. Strain MC, Little SJ, Daar ES, et al. Effect of treatment, during primary infection, on establishment and clearance of cellular reservoirs of HIV-1. *J Infect Dis.* 2005;191(9):1410-1418.
186. Ananworanich J, Schuetz A, Vandergeeten C, et al. Impact of multi-targeted antiretroviral treatment on gut T cell depletion and HIV reservoir seeding during acute HIV infection. *PLoS One.* 2012;7(3):e33948.
187. Josefsson L, von Stockenstrom S, Faria NR, et al. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci U S A.* 2013;110(51):E4987-4996.
188. Pinzone MR, VanBelzen DJ, Weissman S, et al. Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by clonal expansion. *Nat Commun.* 2019;10(1):728.
189. Gandhi RT CJ, Bosch R, et al. Intact proviral DNA levels decline in people with HIV on antiretroviral therapy. In. Conference on Retroviruses and Opportunistic Infections (CROI). March 8-11, 2020. Boston2020.
190. Wang Z, Gurule EE, Brennan TP, et al. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proc Natl Acad Sci U S A.* 2018;115(11):E2575-E2584.
191. Chiappini E, Galli L, Tovo PA, et al. Virologic, immunologic, and clinical benefits from early combined antiretroviral therapy in infants with perinatal HIV-1 infection. *AIDS.* 2006;20(2):207-215.
192. Shearer WT, Quinn TC, LaRussa P, et al. Viral load and disease progression in infants infected with human immunodeficiency virus type 1. *New Engl J Med.* 1997;336(19):1337-1342.
193. Rainwater-Lovett K, Ziemniak C, Watson D, et al. Paucity of Intact Non-Induced Provirus with Early, Long-Term Antiretroviral Therapy of Perinatal HIV Infection. *PLoS One.* 2017;12(2):e0170548.
194. Dhummakupt A, Siems LV, Singh D, et al. The Latent Human Immunodeficiency Virus (HIV) Reservoir Resides Primarily in CD32-CD4+ T Cells in Perinatally HIV-Infected Adolescents With Long-Term Virologic Suppression. *J Infect Dis.* 2019;219(1):80-88.
195. Katusiime MG, Halvas EK, Wright I, et al. Intact HIV Proviruses Persist in Children Seven to Nine Years after Initiation of Antiretroviral Therapy in the First Year of Life. *J Virol.* 2020;94(4).
196. Luzuriaga K, Tabak B, Garber M, et al. HIV type 1 (HIV-1) proviral reservoirs decay continuously under sustained virologic control in HIV-1-infected children who received early treatment. *J Infect Dis.* 2014;210(10):1529-1538.

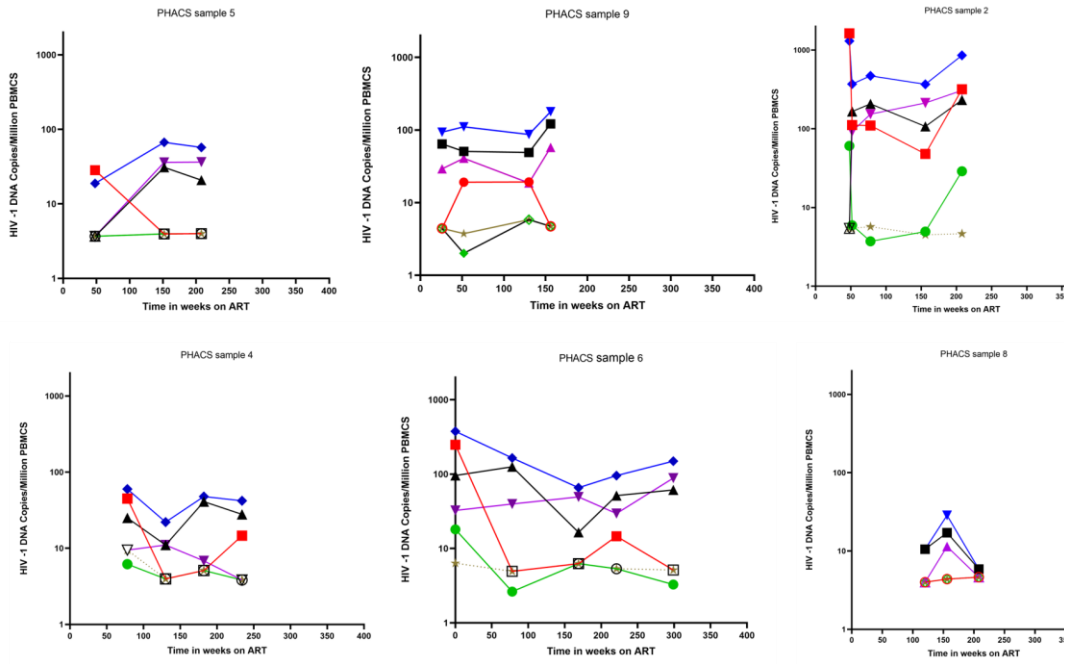
197. Wirotpaisankul P, Lapphra K, Maleesatharn A, et al. HIV seronegativity in children, adolescents and young adults living with perinatally acquired HIV: A cross-sectional study in Thailand. *J Int AIDS Soc.* 2020;23(9):e25614.
198. Brumme ZL, Sudderuddin H, Ziemniak C, et al. Genetic complexity in the replication-competent latent HIV reservoir increases with untreated infection duration in infected youth. *AIDS.* 2019;33(2):211-218.
199. Martinez-Bonet M, Puertas MC, Fortuny C, et al. Establishment and Replenishment of the Viral Reservoir in Perinatally HIV-1-infected Children Initiating Very Early Antiretroviral Therapy. *Clin Infect Dis.* 2015;61(7):1169-1178.
200. Faye A, Le Chenadec J, Dollfus C, et al. Early versus deferred antiretroviral multidrug therapy in infants infected with HIV type 1. *Clin Infect Dis.* 2004;39(11):1692-1698.
201. Zanchetta M, Anselmi A, Vendrame D, et al. Early therapy in HIV-1-infected children: effect on HIV-1 dynamics and HIV-1-specific immune response. *Antivir Ther.* 2008;13(1):47-55.
202. McManus M, Mick E, Hudson R, et al. Early Combination Antiretroviral Therapy Limits Exposure to HIV-1 Replication and Cell-Associated HIV-1 DNA Levels in Infants. *PLoS One.* 2016;11(4):e0154391.
203. Violari A, Cotton MF, Gibb DM, et al. Early antiretroviral therapy and mortality among HIV-infected infants. *N Engl J Med.* 2008;359(21):2233-2244.
204. Dhummakupt A, Rubens JH, Anderson T, et al. Differences in inducibility of the latent HIV reservoir in perinatal and adult infection. *JCI Insight.* 2020;5(4).
205. Uprety P, Chadwick EG, Rainwater-Lovett K, et al. Cell-Associated HIV-1 DNA and RNA Decay Dynamics During Early Combination Antiretroviral Therapy in HIV-1-Infected Infants. *Clin Infect Dis.* 2015;61(12):1862-1870.
206. Freguja R, Gianesin K, Mosconi I, et al. Regulatory T cells and chronic immune activation in human immunodeficiency virus 1 (HIV-1)-infected children. *Clin Exp Immunol.* 2011;164(3):373-380.
207. Ananworanich J, Puthanakit T, Suntarattiwong P, et al. Reduced markers of HIV persistence and restricted HIV-specific immune responses after early antiretroviral therapy in children. *AIDS.* 2014;28(7):1015-1020.
208. Saitoh A, Hsia K, Fenton T, et al. Persistence of human immunodeficiency virus (HIV) type 1 DNA in peripheral blood despite prolonged suppression of plasma HIV-1 RNA in children. *J Infect Dis.* 2002;185(10):1409-1416.
209. Van Zyl GU, Katusiime MG, Wiegand A, et al. No evidence of HIV replication in children on antiretroviral therapy. *J Clin Invest.* 2017;127(10):3827-3834.
210. Patel K, Hernan MA, Williams PL, et al. Long-term effects of highly active antiretroviral therapy on CD4+ cell evolution among children and adolescents infected with HIV: 5 years and counting. *Clin Infect Dis.* 2008;46(11):1751-1760.
211. Zanchetta M, Walker S, Burighel N, et al. Long-term decay of the HIV-1 reservoir in HIV-1-infected children treated with highly active antiretroviral therapy. *J Infect Dis.* 2006;193(12):1718-1727.
212. Freguja R, Bamford A, Zanchetta M, et al. Long-term clinical, virological and immunological outcomes following planned treatment interruption in HIV-infected children. *HIV Med.* 2020.
213. Klein N, Sefe D, Mosconi I, et al. The immunological and virological consequences of planned treatment interruptions in children with HIV infection. *PLoS One.* 2013;8(10):e76582.
214. Goetghebuer T, Haelterman E, Le Chenadec J, et al. Effect of early antiretroviral therapy on the risk of AIDS/death in HIV-infected infants. *AIDS.* 2009;23(5):597-604.
215. Cotton MF, Violari A, Otwombe K, et al. Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: results from the

children with HIV early antiretroviral (CHER) randomised trial. *Lancet.* 2013;382(9904):1555-1563.
216. van Zyl GU, Bedison MA, van Rensburg AJ, Laughton B, Cotton MF, Mellors JW. Early Antiretroviral Therapy in South African Children Reduces HIV-1-Infected Cells and Cell-Associated HIV-1 RNA in Blood Mononuclear Cells. *J Infect Dis.* 2015;212(1):39-43.
217. Luzuriaga K, Wu H, McManus M, et al. Dynamics of human immunodeficiency virus type 1 replication in vertically infected infants. *J Virol.* 1999;73(1):362-367.
218. De Rossi A, Walker AS, De Forni D, Gibb DM, Paediatric European Network for Treatment of A. Biphasic decay of cell-associated HIV-1 DNA in HIV-1-infected children on antiretroviral therapy. *AIDS.* 2002;16(14):1961-1963.
219. Garcia-Broncano P, Maddali S, Einkauf KB, et al. Early antiretroviral therapy in neonates with HIV-1 infection restricts viral reservoir size and induces a distinct innate immune profile. *Sci Transl Med.* 2019;11(520).
220. Veldsman KA, Janse van Rensburg A, Isaacs S, et al. HIV-1 DNA decay is faster in children who initiate ART shortly after birth than later. *J Int AIDS Soc.* 2019;22(8):e25368.
221. Organisation WH. Antiretroviral therapy for HIV infection in infants and children. https://www.who.int/hiv/pub/paediatric/infants2010/en/. Published 2010. Accessed 05/03/2021.
222. Bitnun A, Samson L, Chun TW, et al. Early initiation of combination antiretroviral therapy in HIV-1-infected newborns can achieve sustained virologic suppression with low frequency of CD4+ T cells carrying HIV in peripheral blood. *Clin Infect Dis.* 2014;59(7):1012-1019.
223. Martin PJ, Hansen JA, Siadak AW, Nowinski RC. Monoclonal antibodies recognizing normal human T lymphocytes and malignant human B lymphocytes: a comparative study. *J Immunol.* 1981;127(5):1920-1923.
224. Jordan A, Bisgrove D, Verdin E. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *Embo j.* 2003;22(8):1868-1877.
225. Symons J, Chopra A, Malatinkova E, et al. HIV integration sites in latently infected cell lines: evidence of ongoing replication. *Retrovirology.* 2017;14(1):2.
226. Josefsson L, Palmer S, Faria NR, et al. Single cell analysis of lymph node tissue from HIV-1 infected patients reveals that the majority of CD4+ T-cells contain one HIV-1 DNA molecule. *PLoS Pathog.* 2013;9(6):e1003432.
227. Luzuriaga K, McManus M, Catalina M, et al. Early therapy of vertical human immunodeficiency virus type 1 (HIV-1) infection: control of viral replication and absence of persistent HIV-1-specific immune responses. *J Virol.* 2000;74(15):6984-6991.
228. Klein N, Palma P, Luzuriaga K, et al. Early antiretroviral therapy in children perinatally infected with HIV: a unique opportunity to implement immunotherapeutic approaches to prolong viral remission. *The Lancet Infectious Diseases.* 2015;15(9):1108-1114.
229. Research Blood Drawing Guidelines. https://www.childrenshospital.org/~/media/research-and-innovation/office-of-clinical-investigation/113-research-blood-drawing-guidelines102815.ashx?la=en. Updated 5/1/15. Accessed 04/05/2021.
230. Blood Collection Babies and Children. https://lug.hfhs.org/babiesKids.html. Accessed 04/05/2021.
231. Appropriate Maximum Phlebotomy Volumes. https://www.childrensmn.org/departments/lab/pdf/phleb2.pdf. Accessed 04/05/2021.

232. Simonetti FR, Zhang H, Soroosh GP, et al. Antigen-driven clonal selection shapes the persistence of HIV-1-infected CD4+ T cells in vivo. *J Clin Invest.* 2021;131(3).
233. Butler SL, Johnson EP, Bushman FD. Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. *J Virol.* 2002;76(8):3739-3747.
234. Pierson TC, Kieffer TL, Ruff CT, Buck C, Gange SJ, Siliciano RF. Intrinsic stability of episomal circles formed during human immunodeficiency virus type 1 replication. *J Virol.* 2002;76(8):4138-4144.
235. Bushman F. Measuring covert HIV replication during HAART: the abundance of 2-LTR circles is not a reliable marker. *AIDS.* 2003;17(5):749-750.
236. Whitney JB, Hill AL, Sanisetty S, et al. Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature.* 2014;512(7512):74-77.
237. Vansant G, Chen HC, Zorita E, et al. The chromatin landscape at the HIV-1 provirus integration site determines viral expression. *Nucleic Acids Res.* 2020;48(14):7801-7817.
238. Kint S, Trypsteen W, De Spiegelaere W, et al. Underestimated effect of intragenic HIV-1 DNA methylation on viral transcription in infected individuals. *Clin Epigenetics.* 2020;12(1):36.

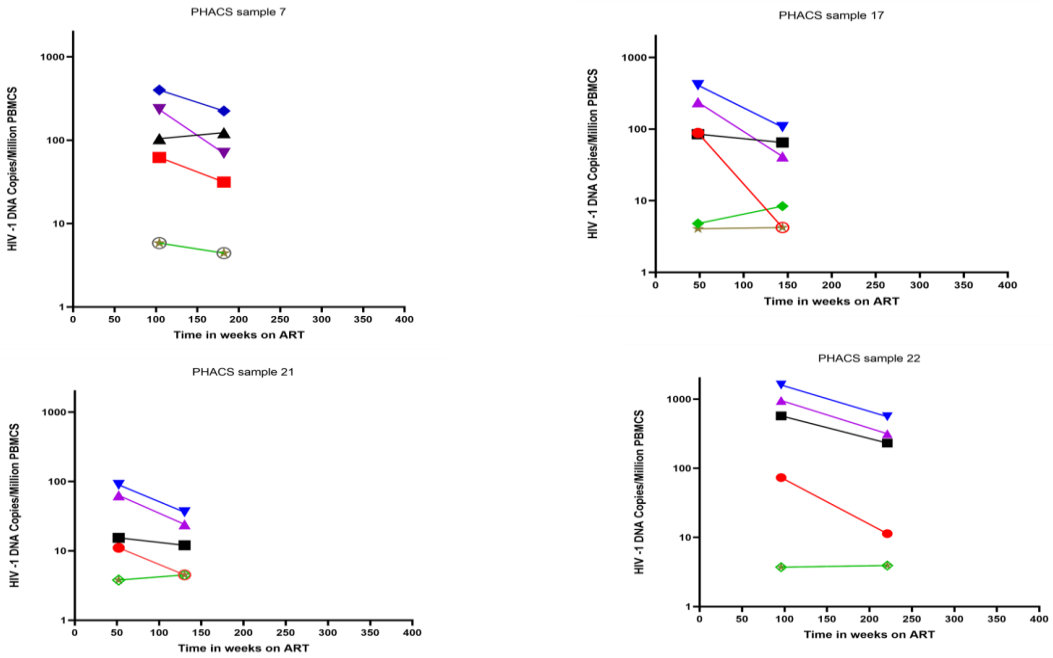
Appendix

The figures below are representative of the individual plots after the analysis of the IPDA results of the PHACS-AMP cohort participants eligible for this study.



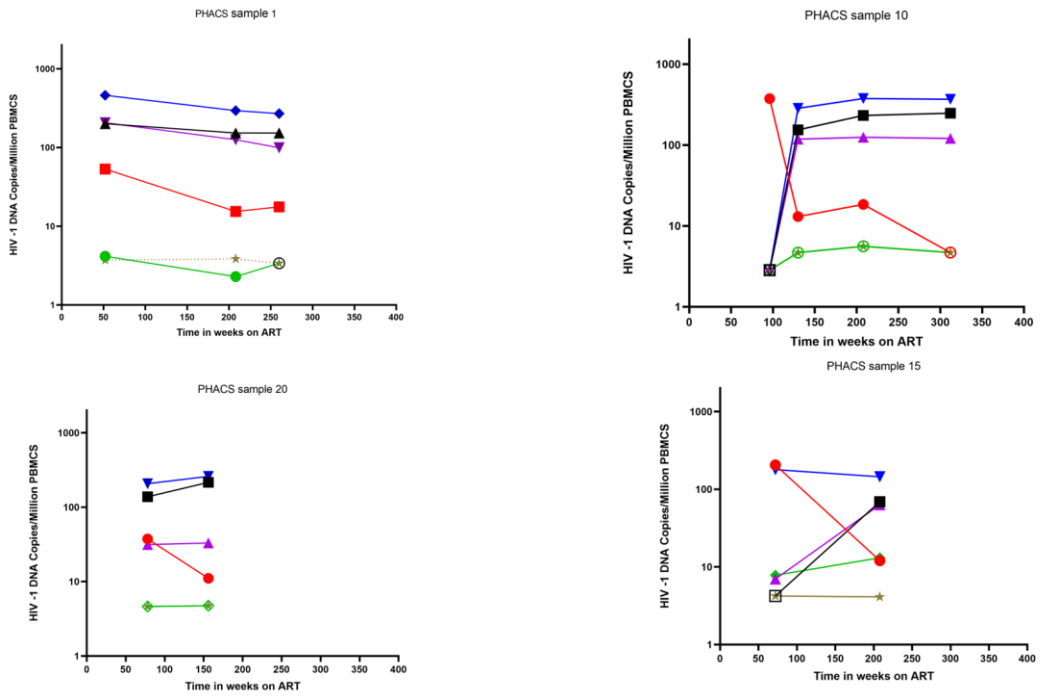
Supplementary Figure 1: Individual plots for participants whose total HIV-1 DNA copies/million PBMCs were in flux over time on ART.

The figures here represent the participants whose Total HIV-1 DNA copies/million PBMCs were in flux over time on ART as a result of flux in intact and defective HIV-1 DNA copies/million PBMCs. The intact HIV-1 DNA copies/million PBMCs are shown in red, the 3' deleted/hypermutated and 5' deleted copies/million PBMCS are shown in black and purple respectively. The total HIV-1 DNA copies/ million PBMCs are shown in blue and the 2LTR copies/million PBMCs are shown in green. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. These samples were set to their respective limit of detection values (shown in gold), which vary with the number of cells analyzed.



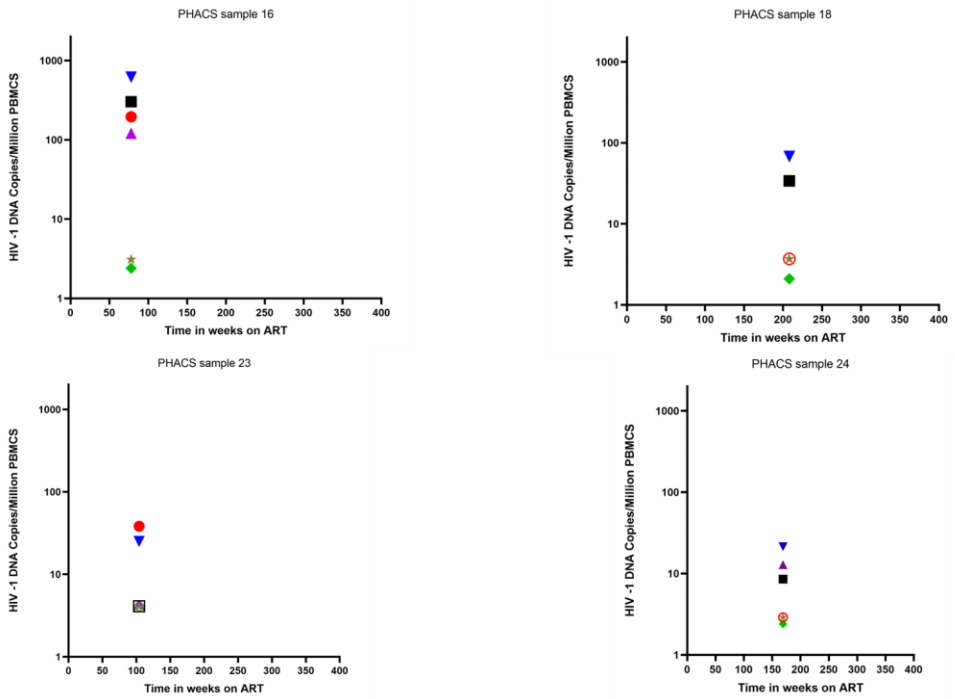
Supplementary Figure 2: Individual plots for participants whose total HIV-1 DNA declined over time on ART.

The figures here represent the participants that whose total HIV-1 DNA declined over time. The intact HIV-1 DNA copies/million PBMCs are shown in red, the 3' deleted/hypermutated and 5' deleted copies/million PBMCs are shown in black and purple respectively. The total HIV-1 DNA copies/ million PBMCs are shown in blue and the 2LTR copies/million PBMCs are shown in green. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. These samples were set to their respective limit of detection values (shown in gold), which vary with the number of cells analyzed.



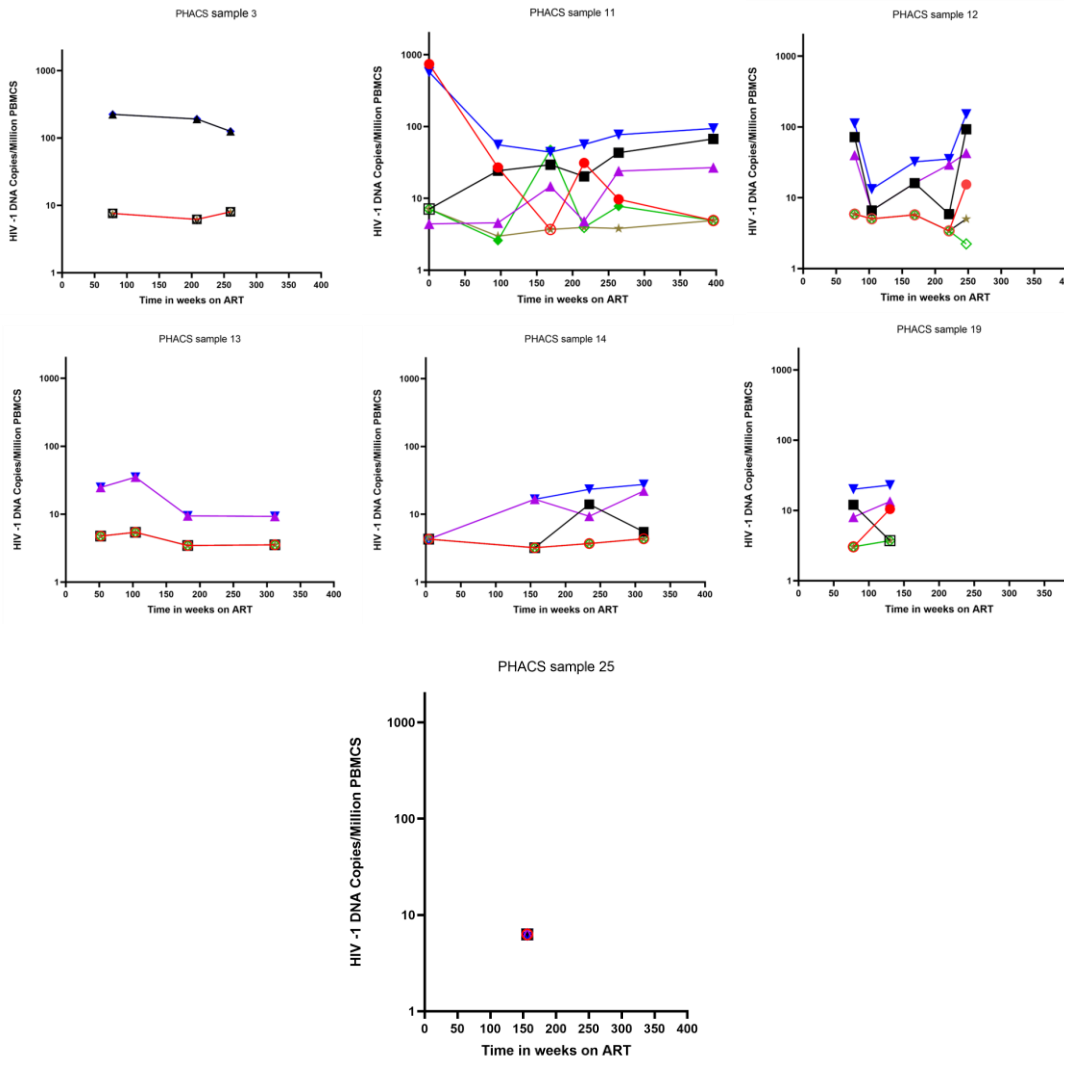
Supplementary Figure 3: Individual plots for participants showing stable total HIV-1 DNA copies/million PBMCs over time on ART.

The figures here represent the participants whose Total HIV-1 DNA copies/million PBMCs were stable over time. The different species of HIV-1 DNA were in a state of flux even though the total HIV-1 DNA was stable. The intact HIV-1 DNA copies/million PBMCs are shown in red, the 3' deleted and 5' deleted copies/million PBMCs are shown in black and purple respectively. The total HIV-1 DNA copies/ million PBMCs are shown in blue and the 2LTR copies/million PBMCs are shown in green. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. These samples were set to their respective limit of detection values (shown in gold), which vary with the number of cells analyzed.



Supplementary Figure 4: Individual plots for participants with only one time point.

The figures here represent the participants that only had one time point eligible for this study. The intact HIV-1 DNA copies/million PBMCs are shown in red, the 3' deleted and 5' deleted copies/million PBMCs are shown in black and purple respectively. The total HIV-1 DNA copies/ million PBMCs are shown in blue and the 2LTR copies/million PBMCs are shown in green. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. These samples were set to their respective limit of detection values (shown in gold), which vary with the number of cells analyzed.



Supplementary Figure 5: Individual plots of participants that need further testing.

The figures here represent the participants whose samples require further testing. The intact HIV-1 DNA copies/million PBMCs are shown in red, the 3' deleted and 5' deleted copies/million PBMCs are shown in black and purple respectively. The total HIV-1 DNA copies/ million PBMCs are shown in blue and the 2LTR copies/million PBMCs are shown in green. The open symbols indicate samples in which HIV-1 DNA was not detected by the assay. These samples were set to their respective limit of detection values (shown in gold), which vary with the number of cells analyzed.

Priya Khetan CV

Work E-mail: pkhetan1@jhmi.edu

Personal E-mail: priyakhetan97@gmail.com

Phone No.: +1 443-447-8049

EDUCATION

2019 – Present

Master of Science (ScM) in Molecular Microbiology and Immunology

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

GPA: 3.71/4.00

2015 – 2018

Bachelor of Science (BSc) in Microbiology and Biochemistry (Distinction)

St. Xavier's College, Autonomous, Mumbai

GPA: 3.79/4.00

Awards: Dr. J. V. Bhatt scholarship for 2nd highest in Microbiology in AY 2017-18

Indian Academy of Sciences Summer Research Fellowship in AY 2016-17.

RESEARCH EXPERIENCE

Nov 2019 – Present

Graduate Student Researcher

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

W. Harry Feinstone Department of Molecular Microbiology and Immunology

Dr. Deborah Persaud's Laboratory

Thesis Topic: Dynamics of Intact Proviruses in Perinatal HIV-1 Infections

Supervisor: Dr. Deborah Persaud and Dr. Adit Dhummakupt

- Optimizing the Intact Proviral DNA Assay (multiplex droplet digital PCR) by: Developing appropriate standards using an HIV-1 cell line, determine the best method for DNA isolation and the performance characteristics of the assay.
- Applying the Intact Proviral DNA Assay to Pediatric HIV AIDS Cohort Study (PHACS-AMP) to understand the reservoir dynamics in children and young adults.

Nov. 2018 – July 2019 **Research Intern**

St. Xavier's College, Autonomous Mumbai

Topic: Generation of Shuttle Vectors for Expression in Bacteria & Yeast

Supervisor: Dr. Vishwas Sarangdhar and Dr. Priya Sunderrajan

- Performed genomic DNA & Plasmid Isolation of Escherichia coli (E. coli) and Saccharomyces cerevisiae (S. cerevisiae). The S. cerevisiae was previously isolated from the skin wild fruits and then mutagenized in the laboratory using UV radiation to obtain a Geneticin sensitive strain.
- Performed transformation of E. coli and S. cerevisiae using economical methods.
- Independently developed ways to transform yeast through disrupting the cell wall using enzymes and bile salts, after established procedures did not work on the laboratory strain.

July 2017 – Oct. 2018 **Student Researcher**

St. Xavier's College, Autonomous, Mumbai

Department of Microbiology

Topic: Extraction and Characterization of a Carotenoid like Pigment Produced by Bacteria Isolated from Soil – Third Year Group Project, Honors Project

Supervisor: Mrs. Sangeetha Chavan, Mrs. Miriam Stewart and Dr. Pampi Chakraborty

- Led the research group project initially and then took it up independently.

- Prepared a proposal and presented it to the faculty.
- The methodology described in the proposal and report led to the successful isolation of a pigment producer from soil (Crowded plate method with 1:10 dilution). In a preliminary spot assay, the pigment showed antibacterial activity. Pigment producers have several broad applications in the industries, especially those that produce carotenoids. The nature of the pigment was characterized using Spectroscopy. The pigment was purified by thin layer and column chromatography. The purified pigment was then sent to another laboratory for Fourier transformed infrared spectroscopy.

Jan 2018. – Feb. 2018 **Student Researcher**

St. Xavier's College, Autonomous, Mumbai

Department of Life Sciences and Biochemistry

Topic: Extraction of Amylase from Ipomoea batatas and Determination of the Effect of Divalent Metal Ions on the Crude Amylase Activity – Third Year Group Project

Supervisors: Dr. Nandita Mangalore and Dr. Priya Jadhav

- Led the research group project and prepared a proposal which was presented to the faculty of Life Sciences and Biochemistry.
- The experiment involved isolation of crude enzyme from sweet potato using established protocols, determining the Km and optimum pH of the crude enzyme using the DNSA assay. The effect of metal ions and their salts was determined using the DNSA assay as well. It was possible to determine the effect of the metal ion had on the crude enzyme and the type of inhibition it showed if applicable through Michelis-Mentin Kinetics.

Dec. 2016 – Mar. 2017 **Student Researcher**

St. Xavier's College, Autonomous, Mumbai

Department of Microbiology

Topic: Isolation of Antibiotic Producers from Soil and

its Effect on Antibiotic Resistant Organisms – Second Year Project

Supervisor: Dr. Karuna Gokarn and Ms. Asma Chikte

- Worked independently on the project as well as presented it independently to the faculty.
- The project methodology resulted in the isolation of a potential antibiotic producer from soil (Crowded plate method with 1:10 dilution and Wilkin's Agar overlay method) which showed a minor effect on growth patterns of Mycobacterium smegmatis in preliminary assays (Central Streak technique).

Apr. 2017 – Jun. 2017 **Medical Virology: Summer Research Fellow**

Rajendra Memorial Research Institute of Medical Sciences

(Indian Academy of Sciences

Topic: Diagnosis and Formulation of an antiviral against Influenza Virus A using Silver Nano Particles

Supervisor: Dr. G. C. Sahoo

- Worked independently on the project. The experiment involved performing ELISAs for Influenza A virus, preparation of Silver nanoparticles using Leishmania and real time PCR to determine the effect of the nanoparticles on the virus. Reached the stage of Silver nanoparticle preparation.
- Alongside this, diagnostic virology was also done wherein ELISAs and DNA/RNA isolations for various viruses such as Hepatitis B, C, A E, Adenovirus, Rotavirus were performed.

May 2016 – Jun 2016 **Clinical Intern**

Jaslok Hospital & Research Center, Department of Microbiology

Topic*:* Observing diagnostic testing and serology

Supervisor: Ms. Barkha Rijhwani

- Performed: Gram Staining, Ziehl Neelsen and Kinyoun Staining, primary processing of the sample i.e., Isolation on media, Antibiotic sensitivity

testing

TECHNICAL SKILLS

Communication & software: Scientific reading, writing and content editing; public speaking; Microsoft Office Applications; Bio-rad Quantasoft for droplet digital PCR analysis; Basic R Statistical Software; Designing graphics using Biorender. Languages spoken: English and Hindi.

Bacterial staining: Gram, Ziehl Nielsen and Kinyoun staining.

Culture & aseptic techniques: Bacteria and yeast culturing; bacterial isolation; preliminary bacterial identification using morphological, cultural, & biochemical characteristics; antibiotic testing; antimicrobial testing; cell line maintenance, serial dilution of cells, whole blood processing for PBMCs, freezing & thawing PBMCs & JLAT cells.

Molecular Biology techniques: Bacterial and yeast transformation; PCR & PCR purification (Qiagen); PAGE & agarose gel electrophoresis; genomic DNA & RNA isolation; plasmid isolation; DNA isolation (Qiagen Qi-amp kit and Gentra Puregene Core B kit), single and multiplex Droplet Digital PCR (Bio-Rad), near Full-length PCR for HIV-1 (B).

Biochemistry techniques: Bioanalytical assays: DNSA (with respect to Enzymology) & biuret, column, and thin-layer chromatography,

ACADEMIC SERVICE

Johns Hopkins Bloomberg School of Public Health

Training and teaching experience

1. Johns Hopkins University LGBTQ+ Office - Safezone Training Facilitator since October 2020
    - My role entails providing a safe and inclusive environment for LGBTQ+ community members, training students, faculty, and staff across the university to provide such spaces as well and advocating for the LGBTQ+ community and their cause.
2. Johns Hopkins University LGBTQ+ Office - Safezone Member
3. JHSPH COVID-19 Mental Health Task Force Fall 2020 Workshop Facilitator
    - My role entails hosting workshops for our student peers during the pandemic to check on them as well as guide them to appropriate resources in case of need.
4. Teaching Assistant: Dr. Jennifer Kavran – Cancer Biology for Term 4, AY 2020 -21
5. Teaching Assistant: Brian Simpson - Writing for Results for Term 3, AY 2019-20 and AY 2020-21
6. Teaching Assistant: Brian Simpson – Unleash Your Writing Superpower: Crafting Clear, Concise and Persuasive Prose for Term 2, AY 2020-21
7. Teaching Assistant: Dr. Prakash Srinivasan - Seminars in MMI for Terms 1 & 2, AY 2020-21
8. Teaching Assistant: Dr. Prakash Srinivasan - Seminars in Molecular Microbiology and Immunology for Terms 3 & 4, AY 2019-20
9. Research Assistant: MMI Insectary Core

Facility AY 2020-21

Leadership experience

1. Molecular Microbiology and Immunology (MMI) Student Committee Masters Student Representative 2020-2021
2. Inclusion, Diversity, Anti-racism, Equity (IDARE) Committee, Student Assembly Representative 2020-2021
    - The Bloomberg School of Public health started a diversity, equity and inclusion initiative in the last few years. A new committee called IDARE: Inclusion, Diversity, Anti-racism and equity was formed earlier this year. The role of IDARE is to solicit feedback from the JHSPH community & create a strategic plan which would be implemented by committees such as the Committee on Equity Diversity & Civility (CEDC). My role in IDARE entails soliciting feedback from the student association & presenting it to the committee followed by reviewing the strategic plan. I hosted listening session for students to gain feedback on what they want to see implemented and written in the IDARE strategic plan.
3. Community Engagement and Dissemination Core (CEDC), Student Representative 2020-2021
    - On CEDC, I serve as the student representative.
4. Virtual Teaching and Education Committee for JHSPH AY 2020-21

    Represented masters students at the Virtual Teaching and Education Committee and advocated for extensions of deadlines for students graduating in spring 2021.
5. Return to Research Committee for Molecular Microbiology and Immunology [MMI] AY 2020- 21.

- My role entailed reviewing the laboratory return to research plans for several laboratories against a checklist and noting compliance with the check list. If there was a lack of compliance or any doubts, I had to bring them up with the faculty in the group for discussion.

6. Johns Hopkins Bloomberg School of Public Health Student Assembly: President AY 2020-21
    - The job of the President of the entire Student Association is to: Ensure that all student voices are heard when representing them at a meeting with the School, University, local, state, national, and international agencies and organizations.
    - Sign official documents on behalf of the Student Assembly.
    - Ensure that the interests of all students at the school are preserved and promoted, and advocate for the students when needed.
    - Improve the relationship of students with the administration by collaborating with the admin and advocating for more transparency where possible.
    - Preside over the Student Assembly meetings
    - I advocated to make unconscious bias training and SafeZone training mandatory for students, faculty and staff.
    - My team and I created summer engagement groups for incoming students to promote networking.
7. Johns Hopkins Bloomberg School of Public Health Student Assembly: President Elect AY 2019-20
    - The role of the President-elect is to: Preside over Student Assembly (SA) in absence of the President, work

collaboratively within SA and learn how committees function, serve as the Student Assembly's office manager, assume the duties of other Student Assembly Executive Officers when there are vacancies in those officer positions.

- The President elect can also bring up ideas to ensure that the student body is well represented and served. For eg: I came up with the idea for a school-based day care for our students who are parents + more transparency between SA, Admin and the students.

8. Johns Hopkins Bloomberg School of Public Health Student Assembly: Co-VP Social and Cultural Affairs AY 2019-20
    - The role of the Vice president (VP) of the Social and Cultural Affairs Committee is: To organise events which celebrate the cultural diversity in the school as well as social and networking events targeting the entire Student Association.
    - We are also responsible for providing opportunities for inter-school interactions for the Student Association.
    - In my year as VP, we hosted the following events: Spring Formal, Game nights, Cultural Night, Virtual fitness classes when the lockdown was initiated. Planned therapy dog sessions & interschool Talent Show (cancelled owing to the pandemic).
9. MMI Student committee: Student Welcoming Committee AY 2019-2020

St. Xavier's College Autonomous, Mumbai

Training and teaching experience

1. Student of the Honors Certificate Program from the Department of

Microbiology from 2015-18 and Department of Chemistry from 2016-18.

2. Editor of Content for the Annual Magazine of the Department of Microbiology, ‘The Michronicle’ in AY 2017-18.
    - As part of the Content team, I had to contact people to write articles for the magazine, decide the titles for each section & for articles that did not have them. I also had to edit articles by checking the flow, content, grammar, appropriateness, citations & plagiarism.
    - My other duties included editing interviews, coming up with fillers/Scientific facts, helping the marketing team find sponsors & deciding the layout of the magazine.

Leadership experience

1. Malhar Fest Department Coordinator in AY 2017-18
    - In the annual intercollegiate fest (hosted by my college) Malhar, I was a coordinator of the sub-department Floors in the Venue Management group: Assistance. The job included:
    - Deciding the layout of the floor, holding group discussions & selecting volunteers, discussing the layout of the floor with the volunteers & heads of the events departments, maintaining decorum on the floors on the days of the fest, solving conflicts between participants & volunteers (if any), managing the props of each team, and creating judge blocks for the judges of the events departments & coordinating with the security teams
2. Blood Drive Secretary in AY 2016-17.

- Before the Drive, my role entailed: coordinating with the blood bank regarding their arrival discuss the process & layout of the venue, publicizing the event, selecting volunteers via group discussions, reserving the venue, organizing the refreshments & the informative/fun stalls.
- On the main day, I had to: handle any problem (including donors fainting), ensure that the donors were comfortable, maintain decorum, set up & clean up the venue.

3. Volunteer: Ithaka Fest in AY 2016-1
4. Volunteer for Blood Drive and Malhar Fest in AY 2015-16.

Athletic and Miscellaneous Experiences

1. Volunteer at Ashray, an orphanage for AIDS affected and infected children under Dr. Mamatha Lala in 2015 and as a part of community service in 2016.
    - Ashray is a temporary residential shelter for children infected and affected by HIV, which is run by the Committed Communities Development Trust. Dr. Mamatha Lala works there as a consultant. I volunteered under her & got to observe the clinical aspects such as: her check-ups on the children, prescription of treatments, how she dealt with the families of the children. I had the opportunity to interact with the children. My duties involved helping, the children with their studies & building their communication skills especially, using team sports/activities as a medium.

- When I went back for my college's Social Involvement Program, my work entailed reading stories to the children & managing their physical activity. I also had the opportunity to help them use their imagination through art. For their physical activity, we played a few games involving relay races, soccer warm up exercises & dancing. I ensured that the older girls & boys tried to read the storybooks rather than have me read it to them & held quizzes. For artwork, I gave them crayons, asked them to meditate & draw their happy place.

2. Member of the Football Team in AY 2015-2016 in St. Xavier’s College
3. Football Team player from 2007-2013 in Middle and High School.
4. PADI Advanced Open Water Diver since September 2018.
5. Games Vice Captain and Games Captain from 2011-12 and 2012-13 respectively high school
6. Attended the Advanced Space Academy of NASA at the Huntsville Center in 2012.